```
DE   Guanosine rich oligonucleotide used to treat viral infection.
XX
KW   Guanosine; tetrad; inhibition; replication; virus; treatment;
KW   therapy; infection; herpes simplex virus; human papilloma virus;
KW   Epstein-Barr virus; HIV, adenovirus; respiratory syncitial virus;
KW   hepatitis B virus; human cytomegalovirus; ss.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   misc_feature    26
FT                   /*tag= a
FT                   /mod_base=
FT                   /note= "Amine moiety attached to this base."
FT   misc_feature    1..26
FT                   /*tag= b
FT                   /note= "Phosphorothioate backbone."
XX
PN   WO9425037-A.
XX
PD   10-NOV-1994.
XX
PF   25-APR-1994;    94WO-US04529.
XX
PR   23-APR-1993;    93US-0053027.
PR   28-OCT-1993;    93US-0145704.
XX
PA   (BAYU ) BAYLOR COLLEGE MEDICINE.
PA   (TRIP-) TRIPLEX PHARM CORP.
XX
PI   Fennewald S,  Hogan ME,  Ojwang JO,  Rando RF,  Zendegui JG;
XX
DR   WPI; 1994-357890/44.
XX
PT   Oligo-nucleotide(s) rich in guanosine which form guanosine
PT   tetrads  - used to treat viral infections, e.g. herpes-virus and
PT   HIV
XX
PS   Claim 41; Page 58; 101pp; English.
XX
CC   The oligonucleotides (See AAQ79201-52)  can be used to treat viral
CC   infections. The oligonucleotides inhibit viral replication by
CC   forming guanosine tetrads which form a stabilised 3D structure.
CC   Preferred oligonucleotides contain at least 2 runs of at least 2
CC   guanosine bases and may be capped at the 3' terminus with a modifier
CC   selected from polyamine, poly-L-lysine, cholesterol and
CC   propanolamine.  They may also have a modified phosphodiester linkage
CC   or be modified to contain a phosphorothioate linkage.  They are used
CC   to treat infections with viruses such as herpes simplex virus, human
CC   papilloma virus, Epstein-Barr virus, HIV, adenovirus, respiratory
CC   syncitial virus, hepatitis B virus or human cytomegalovirus.
XX
SQ   Sequence 26 BP; 0 A; 0 C; 17 G; 9 T; 0 other;

  Query Match             68.0%;  Score 13.6;  DB 15;  Length 26;
  Best Local Similarity   80.0%;  Pred. No. 5.7e+03;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 CGACCCCACAGACAGCCCCC 20
           | |||||||  ||| |||||
Db      23 CCACCCCACCAACAACCCCC 4


RESULT 13
AAQ79220/c
ID   AAQ79220 standard; DNA; 26 BP.
XX
AC   AAQ79220;
XX
DT   17-JUL-1995  (first entry)
XX
DE   Guanosine rich oligonucleotide used to treat viral infection.
XX
KW   Guanosine; tetrad; inhibition; replication; virus; treatment;
KW   therapy; infection; herpes simplex virus; human papilloma virus;
KW   Epstein-Barr virus; HIV, adenovirus; respiratory syncitial virus;
KW   hepatitis B virus; human cytomegalovirus; ss.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   misc_feature    26
FT                   /*tag= a
FT                   /mod_base=
FT                   /note= "Amine moiety attached to this base."
XX
PN   WO9425037-A.
XX
PD   10-NOV-1994.
XX
PF   25-APR-1994;    94WO-US04529.
XX
PR   23-APR-1993;    93US-0053027.
PR   28-OCT-1993;    93US-0145704.
XX
PA   (BAYU ) BAYLOR COLLEGE MEDICINE.
PA   (TRIP-) TRIPLEX PHARM CORP.
XX
PI   Fennewald S,  Hogan ME,  Ojwang JO,  Rando RF,  Zendegui JG;
XX
DR   WPI; 1994-357890/44.
XX
PT   Oligo-nucleotide(s) rich in guanosine which form guanosine
PT   tetrads  - used to treat viral infections, e.g. herpes-virus and
PT   HIV
XX
PS   Claim 41; Page 53; 101pp; English.
XX
CC   The oligonucleotides (See AAQ79201-52)  can be used to treat viral
CC   infections. The oligonucleotides inhibit viral replication by
CC   forming guanosine tetrads which form a stabilised 3D structure.
CC   Preferred oligonucleotides contain at least 2 runs of at least 2
CC   guanosine bases and may be capped at the 3' terminus with a modifier
CC   selected from polyamine, poly-L-lysine, cholesterol and
CC   propanolamine.  They may also have a modified phosphodiester linkage
CC   or be modified to contain a phosphorothioate linkage.  They are used
CC   to treat infections with viruses such as herpes simplex virus, human
CC   papilloma virus, Epstein-Barr virus, HIV, adenovirus, respiratory
CC   syncitial virus, hepatitis B virus or human cytomegalovirus.
XX
SQ   Sequence 26 BP; 0 A; 0 C; 17 G; 9 T; 0 other;

  Query Match             68.0%;  Score 13.6;  DB 15;  Length 26;
  Best Local Similarity   80.0%;  Pred. No. 5.7e+03;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 CGACCCCACAGACAGCCCCC 20
           | |||||||  ||| |||||
Db      23 CCACCCCACCAACAACCCCC 4


RESULT 14
AAT51645/c
ID   AAT51645 standard; DNA; 26 BP.
XX
AC   AAT51645;
XX
DT   12-NOV-1997  (first entry)
XX
DE   Viral integrase inhibiting oligonucleotide.
XX
KW   Human immunodeficiency virus; HIV; Epstein Barr virus; EBV;
KW   herpes simplex virus; HSV; human papilloma virus; HPV; adenovirus;
KW   respiratory syncytial virus; RSV; cytomegalovirus; CMV; hepatitis B;
KW   integrase inhibition; guanosine tetrad; ss.
XX
```

```
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   modified_base   26
FT                   /*tag=  a
FT                   /note= "phosphorothioate backbone and amine moiety
FT                          attached to 3' end"
XX
PN   WO9703997-A1.
XX
PD   06-FEB-1997.
XX
PF   17-JUL-1996;   96WO-US11786.
XX
PR   23-APR-1996;   96US-0016271.
PR   19-JUL-1995;   95US-0001505.
PR   23-OCT-1995;   95US-0535168.
PR   19-MAR-1996;   96US-0013688.
PR   25-MAR-1996;   96US-0014007.
PR   17-APR-1996;   96US-0015714.
XX
PA   (ARON-) ARONEX PHARM INC.
XX
PI   Fennewald S,  Hogan ME,  Mazumder A,  Ojwang JO,  Pommier Y;
PI   Rando RF, Zendegui JG;
XX
DR   WPI; 1997-132569/12.
XX
PT   Oligo:nucleotide(s) capable of forming guanosine tetrads - inhibit
PT   viral enzyme responsible for integrating viral nucleic acid into the
PT   host genome
XX
PS   Claim 3; Page 155; 245pp; English.
XX
CC   AAT51619-T51698 are oligonucleotides used to inhibit the production
CC   of viruses within a host cell. The oligonucleotides may form guanosine
CC   tetrads (structures formed of eight hydrogen bonds by coordination of
CC   the four oxygen atoms of guanine with alkali cations believed to bind
CC   to the centre of a quadruplex, and by strong stacking interactions) and
CC   are used to prevent the integration of viral nucleic acid into a host
CC   genome. The oligonucleotides inhibit functioning of the integrase enzyme
CC   and hence prevent viral infection. Viral infections that may be treated
CC   include human immunodeficiency virus (HIV), Epstein Barr virus (EBV),
CC   herpes simplex virus (HSV), human papilloma virus (HPV), adenovirus,
CC   respiratory syncytial virus (RSV), cytomegalovirus (CMV) and hepatitis
CC   B virus (HBV), especially HIV-1 infection.
XX
SQ   Sequence 26 BP; 0 A; 0 C; 17 G; 9 T; 0 other;

  Query Match             68.0%;  Score 13.6;  DB 18;  Length 26;
  Best Local Similarity   80.0%;  Pred. No. 5.7e+03;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 CGACCCCACAGACAGCCCCC 20
           | ||||||| ||| |||||
Db      23 CCACCCCACCAACAACCCCC 4


RESULT 15
AAT51638/c
ID   AAT51638 standard; DNA; 26 BP.
XX
AC   AAT51638;
XX
DT   12-NOV-1997  (first entry)
XX
DE   Viral integrase inhibiting oligonucleotide.
XX
KW   Human immunodeficiency virus; HIV; Epstein Barr virus; EBV;
KW   herpes simplex virus; HSV; human papilloma virus; HPV; adenovirus;
KW   respiratory syncytial virus; RSV; cytomegalovirus; CMV; hepatitis B;
KW   integrase inhibition; guanosine tetrad; ss.
XX
```

```
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   modified_base   26
FT                   /*tag=  a
FT                   /note= "amine moiety attached to 3' end"
XX
PN   WO9703997-A1.
XX
PD   06-FEB-1997.
XX
PF   17-JUL-1996;   96WO-US11786.
XX
PR   23-APR-1996;   96US-0016271.
PR   19-JUL-1995;   95US-0001505.
PR   23-OCT-1995;   95US-0535168.
PR   19-MAR-1996;   96US-0013688.
PR   25-MAR-1996;   96US-0014007.
PR   17-APR-1996;   96US-0015714.
XX
PA   (ARON-) ARONEX PHARM INC.
XX
PI   Fennewald S,  Hogan ME,  Mazumder A,  Ojwang JO,  Pommier Y;
PI   Rando RF, Zendegui JG;
XX
DR   WPI; 1997-132569/12.
XX
PT   Oligo:nucleotide(s) capable of forming guanosine tetrads - inhibit
PT   viral enzyme responsible for integrating viral nucleic acid into the
PT   host genome
XX
PS   Claim 3; Page 150; 245pp; English.
XX
CC   AAT51619-T51698 are oligonucleotides used to inhibit the production
CC   of viruses within a host cell. The oligonucleotides may form guanosine
CC   tetrads (structures formed of eight hydrogen bonds by coordination of
CC   the four oxygen atoms of guanine with alkali cations believed to bind
CC   to the centre of a quadruplex, and by strong stacking interactions) and
CC   are used to prevent the integration of viral nucleic acid into a host
CC   genome. The oligonucleotides inhibit functioning of the integrase enzyme
CC   and hence prevent viral infection. Viral infections that may be treated
CC   include human immunodeficiency virus (HIV), Epstein Barr virus (EBV),
CC   herpes simplex virus (HSV), human papilloma virus (HPV), adenovirus,
CC   respiratory syncytial virus (RSV), cytomegalovirus (CMV) and hepatitis
CC   B virus (HBV), especially HIV-1 infection.
XX
SQ   Sequence 26 BP; 0 A; 0 C; 17 G; 9 T; 0 other;

  Query Match             68.0%;  Score 13.6;  DB 18;  Length 26;
  Best Local Similarity   80.0%;  Pred. No. 5.7e+03;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 CGACCCCACAGACAGCCCCC 20
           | ||||||| ||| |||||
Db      23 CCACCCCACCAACAACCCCC 4


Search completed: November 23, 2002, 06:29:06
Job time : 101.6 secs
```

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 25, 2002, 09:10:06 ; Search time 755.55 Seconds
(without alignments)
428.707 Million cell updates/sec

Title: US-09-296-264-13
Perfect score: 20
Sequence: 1 tctctgtcctccaaatcgaa 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 357874

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	14.8	74.0	23	17	AZ335717	AZ335717 1M0065J20
c	2	14.2	71.0	80	14	BQ090230	BQ090230 rc63b02.y
	3	14	70.0	65	17	AZ862917	AZ862917 2M0170M16
	4	14	70.0	74	13	BM565947	BM565947 rt05e02.y
	5	14	70.0	75	13	BM566451	BM566451 rt05e02.y
c	6	13.8	69.0	25	17	AZ593505	AZ593505 1M0405M01
	7	13.8	69.0	70	9	AI626412	AI626412 fc08d06.y
	8	13.6	68.0	69	13	BJ041332	BJ041332 BJ041332
	9	13.6	68.0	88	9	AA123311	AA123311 mn34b08.r
c	10	13.6	68.0	88	14	T24714	T24714 EST289 Huma
c	11	13.6	68.0	90	9	AI951126	AI951126 wx63h03.x
	12	13.6	68.0	90	13	BJ057629	BJ057629 BJ057629
	13	13.6	68.0	90	17	AL755734	AL755734 Arabidops
c	14	13.6	68.0	97	12	BF012409	BF012409 ux55h04.y
c	15	13.4	67.0	100	17	BH223688	BH223688 1006114D0
c	16	13.2	66.0	45	17	AZ664779	AZ664779 1M0545D22
	17	13.2	66.0	64	10	AW620171	AW620171 SWOvAFCAP
c	18	13.2	66.0	67	9	AA542491	AA542491 fa07b05.r
	19	13.2	66.0	72	17	TA267C10Q	AL486838 T. brucei
c	20	13.2	66.0	77	17	BH856313	BH856313 SALK_0799
	21	13.2	66.0	89	10	AW458487	AW458487 sh09g11.y
c	22	13.2	66.0	93	17	AZ435176	AZ435176 1M0222C04
	23	13.2	66.0	100	9	AA600769	AA600769 np44f05.s
c	24	13.2	66.0	100	17	AZ950344	AZ950344 2M0214F19
	25	13	65.0	62	17	AZ819711	AZ819711 2M0091B05
c	26	12.8	64.0	72	10	AW085889	AW085889 xc60e03.x
c	27	12.8	64.0	85	17	BH231998	BH231998 1006165E0
	28	12.8	64.0	89	10	AV949920	AV949920 AV949920
c	29	12.8	64.0	91	9	AA547468	AA547468 vk69h06.s
	30	12.8	64.0	91	10	AV964495	AV964495 AV964495
c	31	12.8	64.0	98	17	AL764239	AL764239 Arabidops
	32	12.8	64.0	99	10	AV949925	AV949925 AV949925
	33	12.6	63.0	52	17	AZ590017	AZ590017 1M0399D24
	34	12.6	63.0	58	12	BF595671	BF595671 su91a11.y
c	35	12.6	63.0	59	17	AZ829183	AZ829183 2M0106K09
c	36	12.6	63.0	67	17	AZ344052	AZ344052 1M0077C23
c	37	12.6	63.0	71	10	BE306519	BE306519 601102723
	38	12.6	63.0	75	12	BF596709	BF596709 su73f11.y
	39	12.6	63.0	81	14	Z18343	Z18343 HSDHEKL46 S
c	40	12.6	63.0	82	17	BH623251	BH623251 1007091E0
	41	12.6	63.0	84	17	AZ581322	AZ581322 1M0370H01
c	42	12.6	63.0	84	17	BH217381	BH217381 1006054F0
	43	12.6	63.0	86	12	BG511442	BG511442 sad17h06.
	44	12.6	63.0	89	17	HSEXTR35	Y10934 H.sapiens D
c	45	12.6	63.0	90	9	AI893626	AI893626 mm90d05.x

ALIGNMENTS

RESULT 1
AZ335717
LOCUS AZ335717 23 bp DNA linear GSS 29-SEP-2000
DEFINITION 1M0065J20R Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC1M0065J20 R, DNA sequence.
ACCESSION AZ335717
VERSION AZ335717.1 GI:10404306
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 23)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D.,Weiss,R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00

```
            Plate: 0065  row: J  column: 20
            Seq primer: CACACAGGAAACAGCTATGACC
            Class: plasmid ends
            High quality sequence stop: 23.
FEATURES              Location/Qualifiers
     source           1. .23
                      /organism="Mus musculus"
                      /strain="C57BL/6J"
                      /db_xref="taxon:10090"
                      /clone="UUGC1M0065J20"
                      /clone_lib="Mouse 10kb plasmid UUGC1M library"
                      /sex="Male"
                      /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                      /note="Vector: PWD42nv; Purified genomic DNA from M.
                      musculus C57BL/6J (male) was obtained  from the Jackson
                      Laboratory Mouse DNA Resource
                      (http://www.jax.org/resources/documents/dnares/). The DNA
                      was hydrodynamically sheared by repeated passage through a
                      0.005 inch orifice at constant velocity. The sheared DNA
                      was blunt end-repaired with T4 DNA polymerase and T4
                      polynucleotide kinase. Adaptor oligonucleotides were
                      ligated to the blunt ends in high  molar excess. The
                      adaptored DNA was purified and size-selected for a 9.5 to
                      10.5 kb range using preparative agarose gel
                      electrophoresis. Vector DNA was prepared from a derivative
                      of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                      inducible derivative of plasmid R1. The vector was ligated
                      with adaptors complementary to the insert adaptors and
                      purified. The sheared, adaptored mouse DNA was annealed to
                      adaptored vector DNA, and transformed into
                      chemically-competent E. coli XL10-Gold (Stratagene) cells
                      and selected for ampicillin resistance."
BASE COUNT        5 a      10 c       2 g       6 t
ORIGIN

  Query Match             74.0%;  Score 14.8;  DB 17;  Length 23;
  Best Local Similarity   88.9%;  Pred. No. 3.9e+03;
  Matches   16;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       3 TCTGTCCTCCAAATCGAA 20
           |||| |||||||||| ||
Db       4 TCTGCCCTCCAAATCCAA 21


RESULT 2
BQ090230/c
LOCUS       BQ090230                  80 bp    mRNA    linear   EST 05-APR-2002
DEFINITION  rc63b02.y1 Meloidogyne hapla egg pAMP1 v1 Meloidogyne hapla cDNA
            5', mRNA sequence.
ACCESSION   BQ090230
VERSION     BQ090230.1  GI:20064431
KEYWORDS    EST.
SOURCE      Meloidogyne hapla.
  ORGANISM  Meloidogyne hapla
            Eukaryota; Metazoa; Nematoda; Chromadorea; Tylenchida; Tylenchina;
            Tylenchoidea; Heteroderidae; Meloidogyninae; Meloidogyne.
REFERENCE   1  (bases 1 to 80)
  AUTHORS   McCarter,J., Clifton,S., Chiapelli,B., Pape,D., Martin,J., Wylie,T.
            , Dante,M., Marra,M., Hillier,L., Kucaba,T., Theising,B., Bowers,Y.
            , Gibbons,M., Ritter,E., Bennett,J., Franklin,C., Tsagareishvili,R.
            , Ronko,I., Kennedy,S., Maguire,L., Beck,C., Underwood,K., Steptoe
            ,M., Allen,M., Person,B., Swaller,T., Harvey,N., Schurk,R., Kohn,S.
            , Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and
            Wilson,R.
  TITLE     The Washington Univ. Nematode EST Project, 1999
  JOURNAL   Unpublished (1999)
COMMENT     Contact: McCarter JP
            The Washington Univ. Nematode EST Project, 1999
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            The library was constructed by Claire Murphy and Dr. James McCarter
            at Washington University, St. Louis. DNA Sequencing by: Washington
            University Genome Sequencing Center
            Seq primer: -40RP from Gibco.
FEATURES              Location/Qualifiers
     source           1. .80
                      /organism="Meloidogyne hapla"
                      /db_xref="taxon:6305"
                      /clone_lib="Meloidogyne hapla egg pAMP1 v1"
                      /dev_stage="parasitic adult females"
                      /lab_host="DH10B"
                      /note="Vector: pAMP1 (Gibco); Site_1: NotI; Site_2: SalI;
                      The library was constructed by Claire Murphy and Dr. James
                      McCarter at Washington University, St. Louis. The cDNA was
                      made by using Dynabead oligo-dT priming (Dynal). PCR based
                      library using a modified protocol from the SMART PCR cDNA
                      Synthesis Kit from Clontech. Directionally cloned into the
                      UDG sites ofpAMP1."
BASE COUNT       30 a       8 c      15 g      27 t
ORIGIN

  Query Match             71.0%;  Score 14.2;  DB 14;  Length 80;
  Best Local Similarity   84.2%;  Pred. No. 1.1e+04;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 TCTCTGTCCTCCAAATCGA 19
           | ||  |||||||||||||
Db      51 TTTCAATCCTCCAAATCGA 33


RESULT 3
AZ862917
LOCUS       AZ862917                  65 bp    DNA     linear   GSS 21-FEB-2001
DEFINITION  2M0170M16R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC2M0170M16 R, DNA sequence.
ACCESSION   AZ862917
VERSION     AZ862917.1  GI:13060699
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 65)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0170  row: M  column: 16
            Seq primer: CACACAGGAAACAGCTATGACC
            Class: plasmid ends
            High quality sequence stop: 65.
FEATURES              Location/Qualifiers
     source           1. .65
                      /organism="Mus musculus"
                      /strain="C57BL/6J"
                      /db_xref="taxon:10090"
                      /clone="UUGC2M0170M16"
                      /clone_lib="Mouse 10kb plasmid UUGC1M library"
                      /sex="Male"
                      /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                      /note="Vector: PWD42nv; Purified genomic DNA from M.
```

musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptored DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptored mouse DNA was annealed to adaptored vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 21 a 20 c 10 g 14 t
ORIGIN

Query Match 70.0%; Score 14; DB 17; Length 65;
Best Local Similarity 100.0%; Pred. No. 1.3e+04;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       2 CTCTGTCCTCCAAA 15
           ||||||||||||||
Db      32 CTCTGTCCTCCAAA 45
```

RESULT 4
BM565947
LOCUS BM565947 74 bp mRNA linear EST 20-FEB-2002
DEFINITION rt05e02.y1 Pristionchus pacificus mixed stage SL1 TOPO v1 Murphy Chiapelli McCarter Pristionchus pacificus cDNA 5', mRNA sequence.
ACCESSION BM565947
VERSION BM565947.1 GI:18826550
KEYWORDS EST.
SOURCE Pristionchus pacificus.
ORGANISM Pristionchus pacificus
Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida; Neodiplogasteridae; Pristionchus.
REFERENCE 1 (bases 1 to 74)
AUTHORS McCarter,J., Clifton,S., Chiapelli,B., Pape,D., Martin,J., Wylie,T., Dante,M., Marra,M., Hillier,L., Kucaba,T., Theising,B., Bowers,Y., Gibbons,M., Ritter,E., Bennett,J., Franklin,C., Tsagareishvili,R., Ronko,I., Kennedy,S., Maguire,L., Beck,C., Underwood,K., Steptoe,M., Allen,M., Person,B., Swaller,T., Harvey,N., Schurk,R., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.
TITLE The Washington Univ. Nematode EST Project, 1999
JOURNAL Unpublished (1999)
COMMENT Contact: McCarter JP
The Washington Univ. Nematode EST Project, 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
The library was constructed by Claire Murphy, Brandi Chiapelli, and Dr. James McCarter at Washington University, St. Louis. DNA Sequencing by: Washington University Genome Sequencing Center
Seq primer: -40RP from Gibco.
FEATURES Location/Qualifiers
source 1. .74
/organism="Pristionchus pacificus"
/db_xref="taxon:54126"
/clone_lib="Pristionchus pacificus mixed stage SL1 TOPO v1 Murphy Chiapelli McCarter"
/dev_stage="Mixed stage"
/lab_host="DH10B"
/note="Vector: pCRII-TOPO (Invitrogen); Site_1: EcoRI; Site_2: EcoRI; The library was constructed by Claire Murphy, Brandi Chiapelli, and Dr. James McCarter at Washington University, St. Louis. Oligo(dT)-SL1 PCR based library. Pristionchus pacificus mixed stage cDNA PCR products of size >400 nucleotides containing SL1 on the 5' end and oligo(dT) on the 3' end were non-directionally cloned into pCRII-TOPO(Invitrogen) following the Topo TA cloning protocol. "
BASE COUNT 13 a 29 c 15 g 17 t
ORIGIN

Query Match 70.0%; Score 14; DB 13; Length 74;
Best Local Similarity 100.0%; Pred. No. 1.4e+04;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       2 CTCTGTCCTCCAAA 15
           ||||||||||||||
Db      19 CTCTGTCCTCCAAA 32
```

RESULT 5
BM566451
LOCUS BM566451 75 bp mRNA linear EST 20-FEB-2002
DEFINITION rt05e02.y2 Pristionchus pacificus mixed stage SL1 TOPO v1 Murphy Chiapelli McCarter Pristionchus pacificus cDNA 5', mRNA sequence.
ACCESSION BM566451
VERSION BM566451.1 GI:18827469
KEYWORDS EST.
SOURCE Pristionchus pacificus.
ORGANISM Pristionchus pacificus
Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida; Neodiplogasteridae; Pristionchus.
REFERENCE 1 (bases 1 to 75)
AUTHORS McCarter,J., Clifton,S., Chiapelli,B., Pape,D., Martin,J., Wylie,T., Dante,M., Marra,M., Hillier,L., Kucaba,T., Theising,B., Bowers,Y., Gibbons,M., Ritter,E., Bennett,J., Franklin,C., Tsagareishvili,R., Ronko,I., Kennedy,S., Maguire,L., Beck,C., Underwood,K., Steptoe,M., Allen,M., Person,B., Swaller,T., Harvey,N., Schurk,R., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.
TITLE The Washington Univ. Nematode EST Project, 1999
JOURNAL Unpublished (1999)
COMMENT Contact: McCarter JP
The Washington Univ. Nematode EST Project, 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
The library was constructed by Claire Murphy, Brandi Chiapelli, and Dr. James McCarter at Washington University, St. Louis. DNA Sequencing by: Washington University Genome Sequencing Center
Seq primer: -40RP from Gibco.
FEATURES Location/Qualifiers
source 1. .75
/organism="Pristionchus pacificus"
/db_xref="taxon:54126"
/clone_lib="Pristionchus pacificus mixed stage SL1 TOPO v1 Murphy Chiapelli McCarter"
/dev_stage="Mixed stage"
/lab_host="DH10B"
/note="Vector: pCRII-TOPO (Invitrogen); Site_1: EcoRI; Site_2: EcoRI; The library was constructed by Claire Murphy, Brandi Chiapelli, and Dr. James McCarter at Washington University, St. Louis. Oligo(dT)-SL1 PCR based library. Pristionchus pacificus mixed stage cDNA PCR products of size >400 nucleotides containing SL1 on the 5' end and oligo(dT) on the 3' end were non-directionally cloned into pCRII-TOPO(Invitrogen) following the Topo TA cloning protocol. "
BASE COUNT 14 a 29 c 16 g 16 t
ORIGIN

Query Match 70.0%; Score 14; DB 13; Length 75;

```
  Best Local Similarity   100.0%;  Pred. No. 1.4e+04;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 CTCTGTCCTCCAAA 15
           ||||||||||||||
Db      20 CTCTGTCCTCCAAA 33


RESULT 6
AZ593505/c
LOCUS       AZ593505                 25 bp    DNA     linear   GSS 13-DEC-2000
DEFINITION  1M0405M01F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0405M01 F, DNA sequence.
ACCESSION   AZ593505
VERSION     AZ593505.1  GI:11715695
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 25)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0405  row: M  column: 01
            Seq primer: CGTTGTAAAACGACGGCCAGT
            Class: plasmid ends
            High quality sequence stop: 25.
FEATURES             Location/Qualifiers
     source          1. .25
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC1M0405M01"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
BASE COUNT        8 a      0 c     12 g      5 t
ORIGIN

  Query Match             69.0%;  Score 13.8;  DB 17;  Length 25;

  Best Local Similarity   88.2%;  Pred. No. 1.2e+04;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       4 CTGTCCTCCAAATCGAA 20
           || |||||||||| ||
Db      17 CTCTCCTCCAAATCCAA 1


RESULT 7
AI626412
LOCUS       AI626412                 70 bp    mRNA    linear   EST 07-JUN-2001
DEFINITION  fc08d06.y1 Zebrafish WashU MPIMG EST Danio rerio cDNA clone
            IMAGE:3720779 5', mRNA sequence.
ACCESSION   AI626412
VERSION     AI626412.1  GI:4663212
KEYWORDS    EST.
SOURCE      zebrafish.
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes
            ; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 70)
  AUTHORS   Clark,M., Johnson,S.L., Lehrach,H., Lee,R., Li,F., Marra,M., Eddy
            ,S., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood
            ,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B.,
            Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter,E.,
            Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R.
            and Wilson,R.
  TITLE     WashU Zebrafish EST Project 1998
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Stephen L. Johnson
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: zbrafish@watson.wustl.edu
            cDNA Library Preparation: Matthew Clark. cDNA Library Arrayed by:
            Matthew Clark. DNA Sequencing by: Washington University Genome
            Sequencing Center Clone distribution: Genome Systems, St. Louis,
            Missouri (web address: www.genomesystems.com) (email contact:
            info@genomesystems.com) and Research Genetics, Huntsville, Alabama
            (web address: www.resgen.com) (email contact: info@resgen.com) and
            RessourcenZentrumPrimarDatenbank, Berlin, Germany (web address:
            www.rzpd.de)
            Seq primer: T3 ET from Amersham
            High quality sequence stop: 61
            POLYA=No.
FEATURES             Location/Qualifiers
     source          1. .70
                     /organism="Danio rerio"
                     /db_xref="taxon:7955"
                     /clone="IMAGE:3720779"
                     /clone_lib="Zebrafish WashU MPIMG EST"
                     /sex="mixed"
                     /tissue_type="26 somite embryos, adult livers, shield
                     stage embryos"
                     /lab_host="XL1-blue MRF"
                     /note="Vector: pSPORT1; Site_1: NotI; Site_2: SalI; 1st
                     strand cDNA was primed with a Not I - oligo(dT)15 primer
                     [5'pGACTAGTTCTAGATCGCGAGCGGCCGCCCTTTTTTTTTTTTTTT3'];
                     double-stranded cDNA was ligated to Sal I adaptors (BRL),
                     digested with Not I and cloned into the Not I and Sal I
                     sites of the pSPORT1 vector (BRL). Library was constructed
                     by Matthew Clark (Lehrach lab; ICRF, London and Max Planck
                     Institut fuer Molekulare Genetik,Berlin). cDNAs for EST
                     analysis were selected following oligonucleotide
                     hybridization fingerprinting of arrayed clones from
                     zebrafish late somitogenesis (26 ss), adult liver or
                     embryonic shield stage (5.6 h) libraries.  Fingerprint
                     data were used to computationally cluster cDNAs, and a
                     single cDNA from each cluster was chosen for sequencing.
                     In some cases multiple members of the same cluster were
                     sequenced to assess clustering parameters or single clones
```

were sequenced additional times to assess quality control."
BASE COUNT 22 a 13 c 10 g 24 t 1 others
ORIGIN

Query Match 69.0%; Score 13.8; DB 9; Length 70;
Best Local Similarity 88.2%; Pred. No. 1.7e+04;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 TCTCTGTCCTCCAAATC 17
|||||||||| |||||
Db 9 TCTCTGTCCTAAAAATC 25

RESULT 8
BJ041332
LOCUS BJ041332 69 bp mRNA linear EST 07-DEC-2001
DEFINITION BJ041332 NIBB Mochii normalized Xenopus neurula library Xenopus laevis cDNA clone XL008c10 5', mRNA sequence.
ACCESSION BJ041332
VERSION BJ041332.1 GI:17412158
KEYWORDS EST.
SOURCE African clawed frog.
ORGANISM Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae; Xenopodinae; Xenopus.
REFERENCE 1 (bases 1 to 69)
AUTHORS Kitayama,A., Terasaka,C., Mochii,M., Ueno,N., Shin-i,T. and Kohara,Y.
TITLE Expressed genes in X. laevis embryo
JOURNAL Unpublished (2001)
COMMENT Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
FEATURES Location/Qualifiers
source 1. .69
/organism="Xenopus laevis"
/db_xref="taxon:8355"
/clone="XL008c10"
/clone_lib="NIBB Mochii normalized Xenopus neurula library"
/tissue_type="whole embryo"
/dev_stage="stage 15"
/note="Vector: pBSRN3; Site_1: NotI; Site_2: EcoRI; cDNAs were oligo-dT primed and directionally cloned. Staging according to Nieuwkoop and Faber. Library is subtracted and was constructed by N. Garrett and A.M. Zorn, (Wellcome/CRC Institute). "
BASE COUNT 21 a 13 c 17 g 18 t
ORIGIN

Query Match 68.0%; Score 13.6; DB 13; Length 69;
Best Local Similarity 80.0%; Pred. No. 2.1e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 TCTCTGTCCTCCAAATCGAA 20
||| ||||| ||||||||
Db 42 TCTGTGTCCGTGAAATCGAA 61

RESULT 9
AA123311
LOCUS AA123311 88 bp mRNA linear EST 21-NOV-1996
DEFINITION mn34b08.r1 Beddington mouse embryonic region Mus musculus cDNA clone IMAGE:539799 5' similar to TR:G563368 G563368 CAP-BINDING PROTEIN ;, mRNA sequence.
ACCESSION AA123311
VERSION AA123311.1 GI:1680863
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 88)
AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.
TITLE The WashU-HHMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Marra M/Mouse EST Project
WashU-HHMI Mouse EST Project
Washington University School of MedicineP
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:326735
Possible reversed clone: similarity on wrong strand
Seq primer: -40m13 ET
High quality sequence stop: 35.
FEATURES Location/Qualifiers
source 1. .88
/organism="Mus musculus"
/strain="C57BL/6 x DBA"
/db_xref="taxon:10090"
/clone="IMAGE:539799"
/clone_lib="Beddington mouse embryonic region"
/sex="pooled"
/tissue_type="embryo"
/dev_stage="7.5dpc"
/lab_host="DH12S"
/note="Organ: whole embryo; Vector: pCMV-SPORT; Site_1: SalI; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT. Gastrulating embryos were collected at 7.5dpc from C57BL6 x DBA matings, excluding embryos that had developed head folds and all extraembryonic tissues. Average insert size: 1.3 kb (range: 0.5 - 3.0 kb). Referenced in Development 121, 2479-2489 (1995)"
BASE COUNT 21 a 24 c 15 g 28 t
ORIGIN

Query Match 68.0%; Score 13.6; DB 9; Length 88;
Best Local Similarity 80.0%; Pred. No. 2.3e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 TCTCTGTCCTCCAAATCGAA 20
||| |||||| |||| | ||
Db 64 TCTGTGTCCTGCAAACCCAA 83

RESULT 10
T24714/c
LOCUS T24714 88 bp mRNA linear EST 24-AUG-1995
DEFINITION EST289 Human colorectal cancer Homo sapiens cDNA clone 11C3, mRNA sequence.
ACCESSION T24714
VERSION T24714.1 GI:534339
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutel
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo
REFERENCE 1 (bases 1 to 88)
AUTHORS Frigerio,J.-M., Berthezene,P., Garrido,P., Ortiz,E., Barth
Vasseur,S., Sastre,B., Seleznieff,I., Dagorn,J.-C. and Iov
,J.-L.
TITLE Analysis of 2166 clones from a human colorectal cancer cD

[redacted]

```
            by partial sequencing
  JOURNAL   Hum. Mol. Genet. 4, 37-43 (1995)
  MEDLINE   95227175
COMMENT     Contact: Iovanna JL
            U.315 INSERM
            46 Bd de la Gaye, F-13009 Marseille, France.
            Tel: (33) 91 82 03 15
            Fax: (33) 91 26 62 19
            Email: dagorn@arthur.citi2.fr
            This sequence is one of a series obtained by systematic  sequencing
            of a colorectal cancer cDNA library.
            Seq primer: M13 Forward.
FEATURES             Location/Qualifiers
     source          1. .88
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="11C3"
                     /clone_lib="Human colorectal cancer"
                     /lab_host="E. coli NM522"
                     /note="Vector: pT7T3D; Site_1: Eco RI; Site_2: Not I; mRNA
                     was purified from a colorectal tumour of an adult male.
                     cDNA was constructed and cloned into the pT7T3D phagemid
                     following the manufacturer intructions (Pharmacia)."
BASE COUNT       30 a      9 c     26 g     23 t
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 14;  Length 88;
  Best Local Similarity   80.0%;  Pred. No. 2.3e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 TCTCTGTCCTCCAAATCGAA 20
           ||| | ||||| || |||||
Db      73 TCTTTCTCCTCAAACTCGAA 54


RESULT 11
AI951126/c
LOCUS       AI951126                 90 bp    mRNA    linear   EST 06-SEP-1999
DEFINITION  wx63h03.x1 NCI_CGAP_Br18 Homo sapiens cDNA clone IMAGE:2548373 3',
            mRNA sequence.
ACCESSION   AI951126
VERSION     AI951126.1  GI:5743436
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 90)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Chris Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D. cDNA Library Preparation: Life
            Technologies, Inc. cDNA Library Arrayed by: Christa Prange, The
            I.M.A.G.E. Consortium DNA Sequencing by: Washington University
            Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Seq primer: -40UP from Gibco.
FEATURES             Location/Qualifiers
     source          1. .90
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2548373"
                     /clone_lib="NCI_CGAP_Br18"
                     /tissue_type="four pooled high-grade tumors, including two
                     primary tumors and two metastatic to ovary"
                     /lab_host="DH10B"
                     /note="Organ: breast; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally.  Primer: C
                     Library constructed by Life Technologies."
BASE COUNT       25 a      15 c     20 g     30 t
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 9;  Length 90;
  Best Local Similarity   80.0%;  Pred. No. 2.3e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 TCTCTGTCCTCCAAATCGAA 20
           |||||| ||||||||   ||
Db      27 TCTCTGCCCTCCAAAAAAAA 8


RESULT 12
BJ057629
LOCUS       BJ057629                 90 bp    mRNA    linear   EST 10-DEC-2001
DEFINITION  BJ057629 NIBB Mochii normalized Xenopus tailbud library Xenopus
            laevis cDNA clone XL104k07 5', mRNA sequence.
ACCESSION   BJ057629
VERSION     BJ057629.1  GI:17479710
KEYWORDS    EST.
SOURCE      African clawed frog.
  ORGANISM  Xenopus laevis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus.
REFERENCE   1  (bases 1 to 90)
  AUTHORS   Kitayama,A., Terasaka,C., Mochii,M., Ueno,N., Shin-i,T. and Kohara
            ,Y.
  TITLE     Expressed genes in X. laevis embryo
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Tadasu Shin-i
            Center For Genetic Resource Information
            National Institute of Genetics
            1111 Yata, Mishima, Shizuoka 411-8540, Japan
            Tel: 81-559-81-6856
            Fax: 81-559-81-6855
            Email: tshini@genes.nig.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .90
                     /organism="Xenopus laevis"
                     /db_xref="taxon:8355"
                     /clone="XL104k07"
                     /clone_lib="NIBB Mochii normalized Xenopus tailbud
                     library"
                     /tissue_type="whole embryo"
                     /dev_stage="stage 25"
                     /note="Vector: pBSRN3; Site_1: NotI; Site_2: EcoRI; cDNAs
                     were oligo-dT primed and directionally cloned. Staging
                     according to  Nieuwkoop and Faber. Library is subtracted
                     and was constructed by N. Garrett  and A.M. Zorn,
                     (Wellcome/CRC Institute). "
BASE COUNT       24 a      20 c     24 g     21 t      1 others
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 13;  Length 90;
  Best Local Similarity   80.0%;  Pred. No. 2.3e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 TCTCTGTCCTCCAAATCGAA 20
           ||| |||||   ||||||||
Db      63 TCTGTGTCCGTGAAATCGAA 82


RESULT 13
AL755734
LOCUS       AL755734                 90 bp    DNA     linear   GSS 17-JUN-2002
DEFINITION  Arabidopsis thaliana T-DNA flanking sequence GK-101F05-012072,
            genomic survey sequence.
ACCESSION   AL755734
VERSION     AL755734.1  GI:21488232
KEYWORDS    GSS.
```

```
SOURCE      thale cress.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1
  AUTHORS   Strizhov,N., Li,Y., Rosso,M., Viehoever,P., Dekker,K., Saedler,H.
            and Weisshaar,B.
  TITLE     A pipeline for automated high-throughput generation of FSTs
            (flanking sequence tags) from Arabidopsis thaliana T-DNA
            transformed lines
  JOURNAL   Unpublished
REFERENCE   2
  AUTHORS   Rosso,M., Strizhov,N., Li,Y., Reiss,B., Dekker,K. and Weisshaar,B.
  TITLE     A new Arabidopsis thaliana T-DNA mutagenised population (GABI-Kat)
            for flanking sequence tag based reverse genetics
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 90)
  AUTHORS   Li,Y., Rosso,M., Strizhov,N. and Weisshaar,B.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JUN-2002) Weisshaar B., Max-Planck-Institut fuer
            Zuechtungsforschung, Carl-von-Linne-Weg 10, Koeln, 50829, Germany
COMMENT     This sequence is recovered from the left border of the T-DNA. It
            indicates an insertion close to or within gene At4g33260. The
            sequences are generated at the MPI for Plant Breeding Research in
            the context of the GABI-Kat project. GABI-Kat is part of the German
            Plant Genomics program designated 'GABI'. Information on line
            availability can be found at:
            http://www.mpiz-koeln.mpg.de/GABI-Kat/.
FEATURES             Location/Qualifiers
     source          1. .90
                     /organism="Arabidopsis thaliana"
                     /strain="Columbia 0"
                     /db_xref="taxon:3702"
                     /clone="GK-101F05-012072"
                     /clone_lib="Arabidopsis thaliana T-DNA insertion lines"
                     /note="PCR was performed on DNA from Arabidopsis thaliana
                     plants (T1) which were transformed with the T-DNA from
                     vector pAC161. The lines contain one or more T-DNA
                     insertions. The DNA fragment(s) resulting from the PCR
                     were directly sequenced to determine the genomic sequence
                     flanking the insertion. Sequences displaying significant
                     similarity to the A. thalaina nuclear genome sequence were
                     processed for submission. T-DNA derived sequences were
                     removed"
BASE COUNT       29 a     18 c      9 g     32 t      2 others
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 17;  Length 90;
  Best Local Similarity   80.0%;  Pred. No. 2.3e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 TCTCTGTCCTCCAAATCGAA 20
          |||||||| ||| |||  ||
Db     52 TCTCTGTCATCCTAATTTAA 71


RESULT 14
BF012409/c
LOCUS       BF012409                 97 bp    mRNA    linear   EST 06-OCT-2000
DEFINITION  ux55h04.y1 Soares_NKWMD_mandible Mus musculus cDNA clone
            IMAGE:3514231 5' similar to SW:RIR1_MOUSE P07742
            RIBONUCLEOSIDE-DIPHOSPHATE REDUCTASE M1 CHAIN ;, mRNA sequence.
ACCESSION   BF012409
VERSION     BF012409.1  GI:10712684
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 97)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:1397111
            Trace considered overall poor quality
            Seq primer: -40RP from Gibco
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .97
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:3514231"
                     /clone_lib="Soares_NKWMD_mandible"
                     /tissue_type="mandible"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: NotI; Site_2: EcoRI; 1st strand cDNA
                     was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCCTTAATTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. Library
                     went through one round of normalization. Library
                     constructed by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT       30 a     14 c     39 g     14 t
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 12;  Length 97;
  Best Local Similarity   80.0%;  Pred. No. 2.3e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 TCTCTGTCCTCCAAATCGAA 20
          |||||||| ||| |||  ||
Db     75 TCTCTGTTCTCCAAAGAGCA 56


RESULT 15
BH223688/c
LOCUS       BH223688                100 bp    DNA     linear   GSS 08-NOV-2001
DEFINITION  1006114D05.x1 1006 - RescueMu Grid G Zea mays genomic, DNA
            sequence.
ACCESSION   BH223688
VERSION     BH223688.1  GI:16819892
KEYWORDS    GSS.
SOURCE      Zea mays.
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 100)
  AUTHORS   Walbot,V.
  TITLE     Maize genomic sequences found using engineered RescueMu transposon
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Walbot V
            Department of Biological Sciences
            Stanford University
            855 California Ave, Palo Alto, CA 94304, USA
            Tel: 650 723 2227
            Fax: 650 725 8221
            Email: walbot@stanford.edu
            Very probable ligation site found so sequence was trimmed.
            Post-ligation sequence submitted separately.
            Plate: 1006114  row: 11
            Class: transposon-tagged.
FEATURES             Location/Qualifiers
     source          1. .100
                     /organism="Zea mays"
                     /cultivar="mixed background W23/A188/B73"
                     /db_xref="taxon:4577"
                     /clone_lib="1006 - RescueMu Grid G"
```

/tissue_type="leaf"
/dev_stage="adult"
/lab_host="DH10B"
/note="Organ: leaf; Vector: RescueMu (engineered from pBlueScript backbone); Site_1: BamHI; Site_2: BglII; RescueMu is a 4.9 kb, modified maize Mu transposon designed to allow plasmid rescue from total genomic DNA. Mu elements insert preferentially into transcription units. For more information on RescueMu, go to the web site 'www.zmdb.iastate.edu' and follow the links for 'RescueMu.' Grid G was grown at Stanford in 2000. DNA was extracted from leaf punches, double digested using BamHI and BglII, and ligated to form circular plasmids. DH10B cells were transformed and then screened on LB plates with ampicillin."

BASE COUNT 22 a 28 c 30 g 20 t
ORIGIN

Query Match 67.0%; Score 13.4; DB 17; Length 100;
Best Local Similarity 93.3%; Pred. No. 2.9e+04;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy       3 TCTGTCCTCCAAATC 17
           || |||||||||||||
Db      71 TCCGTCCTCCAAATC 57
```

Search completed: November 26, 2002, 04:08:25
Job time : 767.8 secs

```
FEATURES             Location/Qualifiers
     source          1. .30
                     /organism="unknown"
BASE COUNT        6 a      12 c      6 g      6 t
ORIGIN

  Query Match             69.0%;  Score 13.8;  DB 6;  Length 30;
  Best Local Similarity   88.2%;  Pred. No. 2.4e+04;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       3 CTTCCCACCCTGAATGA 19
           ||||||||||| | ||
Db       6 CTTCCCACCCTGCAGGA 22


RESULT 2
AX286587/c
LOCUS       AX286587                 59 bp    DNA     linear   PAT 21-NOV-2001
DEFINITION  Sequence 18 from Patent WO0179510.
ACCESSION   AX286587
VERSION     AX286587.1  GI:17048674
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Rice,J.H. and Stevenson,F.M.
  TITLE     Materials and methods relating to immune responses to fusion
            proteins
  JOURNAL   Patent: WO 0179510-A 18 25-OCT-2001;
            Cancer Research Ventures Limited (GB)
FEATURES             Location/Qualifiers
     source          1. .59
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="Primer"
BASE COUNT       18 a      11 c      19 g      11 t
ORIGIN

  Query Match             69.0%;  Score 13.8;  DB 6;  Length 59;
  Best Local Similarity   88.2%;  Pred. No. 2.3e+04;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 TGCTTCCCACCCTGAAT 17
           ||||||||||| ||||
Db      54 TGCTTCCCACCTGGAAT 38


RESULT 3
E02957/c
LOCUS       E02957                   45 bp    DNA     linear   PAT 29-SEP-1997
DEFINITION  Signal repeat element of mini satellite DNA located upstream of
            human alpha-2 plasmin inhibitor gene.
ACCESSION   E02957
VERSION     E02957.1  GI:2171179
KEYWORDS    JP 1991147799-A/3.
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 45)
  AUTHORS   Hashimoto,T. and Kurose,K.
  TITLE     NOVEL OLIGONUCLEOTIDE PROBE
  JOURNAL   Patent: JP 1991147799-A 3 24-JUN-1991;
            HOECHST JAPAN LTD
COMMENT     OS   Homo sapiens (human)
            PN   JP 1991147799-A/3
            PD   24-JUN-1991
            PF   02-NOV-1989 JP 1989284909
            PI   HASHIMOTO TAMOTSU, KUROSE KOICHI
            PC   C12Q1/68,C07H21/04,C07K13/00,C12N15/11;
            CC   strandedness: Double;
            CC   topology: Linear;
            CC   *source: tissue_type=liver;
            FH   Key             Location/Qualifiers
            FH
            FT   repeat_unit     1. .45
            FT                   /note='Signal repeat element of mini satellite
             FT                   DNA located
            FT                   upstream of human alpha-2 plasmin inhibitor
            FT                   gene'.
FEATURES             Location/Qualifiers
     source          1. .45
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT        8 a       4 c      21 g       5 t       7 others
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 6;  Length 45;
  Best Local Similarity   72.2%;  Pred. No. 3e+04;
  Matches   13;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 TGCTTCCCACCCTGAATG 18
           ::|:|||||||| | ||
Db      23 YSCYTCCCACCCTCACTG 6


RESULT 4
AX165321/c
LOCUS       AX165321                 51 bp    DNA     linear   PAT 22-JUN-2001
DEFINITION  Sequence 516 from Patent WO0138586.
ACCESSION   AX165321
VERSION     AX165321.1  GI:14546150
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 51)
  AUTHORS   Shimkets,R.A. and Leach,M.
  TITLE     Nucleic acids containing single nucleotide polymorphisms and
            methods of use thereof
  JOURNAL   Patent: WO 0138586-A 516 31-MAY-2001;
            Curagen Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .51
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     variation       26
                     /note="single nucleotide polymorphism
                     Accession number cg43258841"
BASE COUNT       11 a      12 c      11 g      17 t
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 6;  Length 51;
  Best Local Similarity   80.0%;  Pred. No. 3e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 TGCTTCCCACCCTGAATGAT 20
           || | | ||||||| |||||
Db      36 TGTTCCACACCCTGGATGAT 17


RESULT 5
AX446547/c
LOCUS       AX446547                 24 bp    DNA     linear   PAT 03-JUL-2002
DEFINITION  Sequence 3002 from Patent WO0216649.
ACCESSION   AX446547
VERSION     AX446547.1  GI:21695446
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Gunderson,K.
  TITLE     Probes and decoder oligonucleotides
```

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 3, 2002, 14:46:40 ; Search time 302.2 Seconds
(without alignments)
1926.063 Million cell updates/sec

Title: US-09-296-264-14
Perfect score: 20
Sequence: 1 tgcttcccaccctgaatgat 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 995600

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		13.8	69.0	30	6	AR028297	AR028297 Sequence
2	c	13.8	69.0	59	6	AX286587	AX286587 Sequence
3	c	13.6	68.0	45	6	E02957	E02957 Signal repe
4	c	13.6	68.0	51	6	AX165321	AX165321 Sequence
5	c	13.4	67.0	24	6	AX446547	AX446547 Sequence
6	c	13.2	66.0	35	6	I09263	I09263 Sequence 30
7	c	13.2	66.0	45	6	E02955	E02955 Single repe
8		13.2	66.0	77	10	MUSCYP45B	K02409 Mouse cytoc
9	c	13.2	66.0	98	6	A42840	A42840 Sequence 17
10	c	13.2	66.0	98	6	I87346	I87346 Sequence 17
11		13	65.0	49	6	AR002632	AR002632 Sequence
12		13	65.0	49	6	AR099682	AR099682 Sequence
13	c	13	65.0	90	10	MMDNDS27	X55245 M.musculus
14	c	12.8	64.0	21	6	AR074263	AR074263 Sequence
15		12.8	64.0	21	6	AR142728	AR142728 Sequence
16	c	12.8	64.0	21	6	AR142731	AR142731 Sequence
17	c	12.8	64.0	21	6	AX032625	AX032625 Sequence
18		12.8	64.0	21	6	I11403	I11403 Sequence 4
19		12.8	64.0	35	6	AR055030	AR055030 Sequence
20		12.8	64.0	35	6	AR156279	AR156279 Sequence
21		12.8	64.0	35	6	AX343203	AX343203 Sequence
22	c	12.8	64.0	48	6	AX221915	AX221915 Sequence
23	c	12.8	64.0	59	6	AX286588	AX286588 Sequence
24		12.8	64.0	75	6	AX076800	AX076800 Sequence
25		12.8	64.0	75	6	AX076802	AX076802 Sequence
26		12.8	64.0	75	6	AX076804	AX076804 Sequence
27		12.8	64.0	75	6	AX076806	AX076806 Sequence
28		12.8	64.0	75	6	AX139703	AX139703 Sequence
29		12.8	64.0	75	6	AX139705	AX139705 Sequence
30		12.8	64.0	75	6	AX139707	AX139707 Sequence
31		12.8	64.0	75	6	AX139709	AX139709 Sequence
32		12.8	64.0	91	11	G42365	G42365 KAM38 Genom
33		12.8	64.0	95	6	AR002350	AR002350 Sequence
34		12.8	64.0	95	6	AR030853	AR030853 Sequence
35		12.8	64.0	95	6	AR068248	AR068248 Sequence
36		12.8	64.0	95	6	I86799	I86799 Sequence 3
37		12.8	64.0	99	3	OBU90832	U90832 Oscheius br
38		12.6	63.0	32	6	AX300122	AX300122 Sequence
39		12.6	63.0	34	6	AX297798	AX297798 Sequence
40		12.6	63.0	34	6	AX297807	AX297807 Sequence
41		12.6	63.0	34	6	AX297816	AX297816 Sequence
42		12.6	63.0	34	6	AX297825	AX297825 Sequence
43		12.6	63.0	34	6	AX297834	AX297834 Sequence
44		12.6	63.0	34	6	AX297843	AX297843 Sequence
45		12.6	63.0	34	6	BD006756	BD006756 Human G p

ALIGNMENTS

RESULT 1
AR028297
LOCUS AR028297 30 bp DNA linear PAT 29-SEP-1999
DEFINITION Sequence 7 from patent US 5858662.
ACCESSION AR028297
VERSION AR028297.1 GI:5940270
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 30)
AUTHORS Keating,M.T. and Morris,C.A.
TITLE Diagnosis of Williams syndrome and Williams syndrome cognitive profile by analysis of the presence or absence of a LIM-kinase gene
JOURNAL Patent: US 5858662-A 7 12-JAN-1999;

```
  JOURNAL   Patent: WO 0216649-A 3002 28-FEB-2002;
            Illumina, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .24
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="Computer Generated Probe Sequence."
BASE COUNT        5 a      6 c      8 g      5 t
ORIGIN

  Query Match             67.0%;  Score 13.4;  DB 6;  Length 24;
  Best Local Similarity   93.3%;  Pred. No. 4e+04;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       5 TCCCACCCTGAATGA 19
           ||||||||||| |||
Db      21 TCCCACCCTGATTGA 7


RESULT 6
I09263/c
LOCUS       I09263                    35 bp    DNA     linear   PAT 02-DEC-1994
DEFINITION  Sequence 30 from Patent WO 8901940.
ACCESSION   I09263
VERSION     I09263.1  GI:588047
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 35)
  AUTHORS   Fisher,R.A., Gilbert,W., Sato,V.L., Flavell,R.A., Maraganore,J.M.
            and Liu,T.R.
  TITLE     DNA SEQUENCES, RECOMBINANT DNA MOLECULES AND PROCESSES FOR
            PRODUCING SOLUBLE T4 PROTEINS
  JOURNAL   Patent: WO 8901940-A 30 09-MAR-1989;
FEATURES             Location/Qualifiers
     source          1. .35
                     /organism="unknown"
BASE COUNT       13 a      5 c     13 g      4 t
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 6;  Length 35;
  Best Local Similarity   83.3%;  Pred. No. 5.1e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       3 CTTCCCACCCTGAATGAT 20
           |||||| ||||| ||| |
Db      29 CTTCCCCCCCTGTATGTT 12


RESULT 7
E02955/c
LOCUS       E02955                    45 bp    DNA     linear   PAT 29-SEP-1997
DEFINITION  Single repeat element of mini satellite DNA located upstream of
            human alpha-2 plasmin inhibitor gene.
ACCESSION   E02955
VERSION     E02955.1  GI:2171177
KEYWORDS    JP 1991147799-A/1.
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 45)
  AUTHORS   Hashimoto,T. and Kurose,K.
  TITLE     NOVEL OLIGONUCLEOTIDE PROBE
  JOURNAL   Patent: JP 1991147799-A 1 24-JUN-1991;
            HOECHST JAPAN LTD
COMMENT     OS   Artificial gene
            OC   Artificial sequence; Genes.
            PN   JP 1991147799-A/1
            PD   24-JUN-1991
            PF   02-NOV-1989 JP 1989284909
            PI   HASHIMOTO TAMOTSU, KUROSE KOICHI
            PC   C12Q1/68,C07H21/04,C07K13/00,C12N15/11;
            CC   strandedness: Single;
            CC   topology: Linear;
            CC   hypothetical: No;
            CC   anti-sense: No;
            FH   Key             Location/Qualifiers
            FH
            FT   repeat_unit     1. .45
            FT                   /note='single repeat element of mini satellite
             FT                   DNA located
            FT                   upstream of human alpha-2 plasmin inhibitor
            FT                   gene'.
FEATURES             Location/Qualifiers
     source          1. .45
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
BASE COUNT       11 a      6 c     23 g      5 t
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 6;  Length 45;
  Best Local Similarity   83.3%;  Pred. No. 5e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 TGCTTCCCACCCTGAATG 18
           | |||||||||||  | ||
Db      23 TCCTTCCCACCCTCACTG 6


RESULT 8
MUSCYP45B
LOCUS       MUSCYP45B                 77 bp    mRNA    linear   ROD 27-APR-1993
DEFINITION  Mouse cytochrome P-450b (phenobarbital-inducible) gene, partial
            exon 7.
ACCESSION   K02409
VERSION     K02409.1  GI:192896
KEYWORDS    cytochrome; cytochrome P450b.
SOURCE      Mouse (DBA/2N) liver, cDNA to mRNA, clone p40.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 77)
  AUTHORS   Stupans,I., Ikeda,T., Kessler,D.J. and Nebert,D.W.
  TITLE     Characterization of a cDNA clone for mouse phenobarbital-inducible
            cytochrome P-450b
  JOURNAL   DNA 3 (2), 129-137 (1984)
  MEDLINE   84207435
   PUBMED   6547088
FEATURES             Location/Qualifiers
     source          1. .77
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
     CDS             <1. .>77
                     /note="cytochrome p-450b"
                     /codon_start=2
                     /protein_id="AAA37510.1"
                     /db_xref="GI:553905"
                     /translation="SHRLPTLDDRSKMPYTDAVIHEIQR"
BASE COUNT       22 a     23 c     15 g     17 t
ORIGIN      41 bp upstream of HpaII site.

  Query Match             66.0%;  Score 13.2;  DB 10;  Length 77;
  Best Local Similarity   83.3%;  Pred. No. 4.8e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 GCTTCCCACCCTGAATGA 19
           ||| ||||||||  ||||
Db      10 GCTACCCACCCTTGATGA 27


RESULT 9
A42840/c
LOCUS       A42840                    98 bp    DNA     linear   PAT 06-MAR-1997
DEFINITION  Sequence 172 from Patent WO9503412.
ACCESSION   A42840
```

```
VERSION     A42840.1  GI:2298289
KEYWORDS    .
SOURCE      Mycobacterium simiae.
  ORGANISM  Mycobacterium simiae
            Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
            Actinomycetales; Corynebacterineae; Mycobacteriaceae;
            Mycobacterium.
REFERENCE   1  (bases 1 to 98)
  AUTHORS   Mabilat,C. and Christen,R.
  TITLE     NUCLEOTIDE FRAGMENT OF MYCOBACTERIAL RIBOSOMAL RNA 23S, PROBES AND
            PRIMERS DERIVED THEREFROM, REAGENT AND METHOD FOR DETECTING SAID
            FRAGMENT
  JOURNAL   Patent: WO 9503412-A 172 02-FEB-1995;
            BIO MERIEUX (FR)
COMMENT     Other publication CA 2145172 950202
            Other publication FR 2709310 950303.
FEATURES             Location/Qualifiers
     source          1. .98
                     /organism="Mycobacterium simiae"
                     /strain="ATCC 25275"
                     /db_xref="taxon:1784"
BASE COUNT       20 a     28 c     29 g     21 t
ORIGIN

  Query Match            66.0%;  Score 13.2;  DB 6;  Length 98;
  Best Local Similarity  83.3%;  Pred. No. 4.8e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 TGCTTCCCACCCTGAATG 18
          |||||| ||||| ||| |
Db     56 TGCTTCGCACCCCGAAGG 39


RESULT 10
I87346/c
LOCUS       I87346                   98 bp    DNA     linear   PAT 10-JUN-1998
DEFINITION  Sequence 172 from patent US 5703217.
ACCESSION   I87346
VERSION     I87346.1  GI:3207064
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 98)
  AUTHORS   Mabilat,C. and Christen,R.
  TITLE     Nucleotide fragment of the 23S ribosomal RNA of mycobacteria,
            derived probes and primers, reagent and detection method
  JOURNAL   Patent: US 5703217-A 172 30-DEC-1997;
FEATURES             Location/Qualifiers
     source          1. .98
                     /organism="unknown"
BASE COUNT       20 a     28 c     29 g     21 t
ORIGIN

  Query Match            66.0%;  Score 13.2;  DB 6;  Length 98;
  Best Local Similarity  83.3%;  Pred. No. 4.8e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 TGCTTCCCACCCTGAATG 18
          |||||| ||||| ||| |
Db     56 TGCTTCGCACCCCGAAGG 39


RESULT 11
AR002632
LOCUS       AR002632                 49 bp    DNA     linear   PAT 04-DEC-1998
DEFINITION  Sequence 21 from patent US 5741899.
ACCESSION   AR002632
VERSION     AR002632.1  GI:3964186
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 49)
  AUTHORS   Capon,D.J., Tian,H., Smith,D.H., Winslow,G.A. and Siekevitz,M.
  TITLE     Chimeric receptors comprising janus kinase for regulating cellular
            pro liferation
  JOURNAL   Patent: US 5741899-A 21 21-APR-1998;
FEATURES             Location/Qualifiers
     source          1. .49
                     /organism="unknown"
BASE COUNT       11 a     18 c     13 g      7 t
ORIGIN

  Query Match            65.0%;  Score 13;  DB 6;  Length 49;
  Best Local Similarity  100.0%;  Pred. No. 6.4e+04;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      4 TTCCCACCCTGAA 16
          |||||||||||||
Db     33 TTCCCACCCTGAA 45


RESULT 12
AR099682
LOCUS       AR099682                 49 bp    DNA     linear   PAT 14-FEB-2001
DEFINITION  Sequence 21 from patent US 6077947.
ACCESSION   AR099682
VERSION     AR099682.1  GI:12809448
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 49)
  AUTHORS   Capon,D.J., Tian,H., Smith,D.H., Winslow,G.A. and Siekevitz,M.
  TITLE     DNA encoding an intracellular chimeric receptor comprising Janus
            kinase
  JOURNAL   Patent: US 6077947-A 21 20-JUN-2000;
FEATURES             Location/Qualifiers
     source          1. .49
                     /organism="unknown"
BASE COUNT       11 a     18 c     13 g      7 t
ORIGIN

  Query Match            65.0%;  Score 13;  DB 6;  Length 49;
  Best Local Similarity  100.0%;  Pred. No. 6.4e+04;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      4 TTCCCACCCTGAA 16
          |||||||||||||
Db     33 TTCCCACCCTGAA 45


RESULT 13
MMDNDS27/c
LOCUS       MMDNDS27                 90 bp    DNA     linear   ROD 27-AUG-1999
DEFINITION  M.musculus microsatellite D0nds27.
ACCESSION   X55245
VERSION     X55245.1  GI:509152
KEYWORDS    microsatellite DNA.
SOURCE      Mus musculus.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 90)
  AUTHORS   Cornall,R.J., Aitman,T.J., Hearne,C.M. and Todd,J.A.
  TITLE     The generation of a library of PCR-analyzed microsatellite variants
            for genetic mapping of the mouse genome
  JOURNAL   Genomics 10 (4), 874-881 (1991)
  MEDLINE   92009923
   PUBMED   1916820
REFERENCE   2  (bases 1 to 90)
  AUTHORS   Cornell,R.C., Aitman,T.J., Hearne,C.M. and Todd,J.A.
  TITLE     The Generation of a Library of PCR-Analysed Microsatellite Variants
            for Genetic Mapping of the Mouse Genome
  JOURNAL   Unpublished
```

```
REFERENCE   3  (bases 1 to 90)
  AUTHORS   Hearne,C.M.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-OCT-1990) C.M. Hearne, UNIVERSITY OF OXFORD, NUFFIELD
            DEPT. SURGERY, LEVEL 6, JOHN RADCLIFFE HOSP., HEADINGTON OXFORD, UK
FEATURES             Location/Qualifiers
     source          1. .90
                     /organism="Mus musculus"
                     /strain="NOD, ssp.domesticus"
                     /db_xref="taxon:10090"
     satellite       1. .90
                     /note="D0nds27"
BASE COUNT       30 a     13 c     29 g     18 t
ORIGIN

  Query Match            65.0%;  Score 13;  DB 10;  Length 90;
  Best Local Similarity  100.0%;  Pred. No. 6.1e+04;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      7 CCACCCTGAATGA 19
          |||||||||||||
Db     86 CCACCCTGAATGA 74


RESULT 14
AR074263/c
LOCUS       AR074263                 21 bp    DNA     linear   PAT 28-AUG-2000
DEFINITION  Sequence 71 from patent US 5952490.
ACCESSION   AR074263
VERSION     AR074263.1  GI:10001018
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 21)
  AUTHORS   Hanecak,R.C., Anderson,K.P., Bennett,C.Frank., Chiang,M.-Y.,
            Brown-Driver,V.L., Ecker,D.J., Vickers,T.A., Wyatt,J.R. and
            Imbach,J.Louis.
  TITLE     Oligonucleotides having a conserved G4 core sequence
  JOURNAL   Patent: US 5952490-A 71 14-SEP-1999;
FEATURES             Location/Qualifiers
     source          1. .21
                     /organism="unknown"
BASE COUNT        3 a      3 c      7 g      8 t
ORIGIN

  Query Match            64.0%;  Score 12.8;  DB 6;  Length 21;
  Best Local Similarity  87.5%;  Pred. No. 8.6e+04;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      4 TTCCCACCCTGAATGA 19
          | ||||||| ||||||
Db     18 TACCCACCCAGAATGA 3


RESULT 15
AR142728
LOCUS       AR142728                 21 bp    DNA     linear   PAT 08-AUG-2001
DEFINITION  Sequence 11 from patent US 6204000.
ACCESSION   AR142728
VERSION     AR142728.1  GI:15104014
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 21)
  AUTHORS   Dong,J.-T., Barrett,J.Carl., Lamb,P.W. and Isaacs,J.T.
  TITLE     Diagnostic methods and gene therapy using reagents derived from the
            human metastasis suppressor gene KAI1
  JOURNAL   Patent: US 6204000-A 11 20-MAR-2001;
FEATURES             Location/Qualifiers
     source          1. .21
                     /organism="unknown"
BASE COUNT        3 a      10 c      3 g      5 t
ORIGIN

  Query Match            64.0%;  Score 12.8;  DB 6;  Length 21;
  Best Local Similarity  87.5%;  Pred. No. 8.6e+04;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      4 TTCCCACCCTGAATGA 19
          | ||||||||||| |||
Db      6 TGCCCACCCTGACTGA 21


Search completed: December  3, 2002, 18:13:30
Job time : 308.2 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 05:52:31 ; Search time 98.55 Seconds
(without alignments)
457.027 Million cell updates/sec

Title: US-09-296-264-14
Perfect score: 20
Sequence: 1 tgcttcccaccctgaatgat 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 2390332

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	20	100.0	20	21	AAZ31444	Human neuropilin m
c	2	20	100.0	60	24	ABN43139	Human spliced tran
	3	13.8	69.0	30	19	AAT99600	Protein kinase LIM
	4	13.8	69.0	34	18	AAT44095	Human G-protein th
c	5	13.6	68.0	45	12	AAQ11829	Probe to minisatel
c	6	13.6	68.0	51	23	ABL00525	Human silent nonco
c	7	13.4	67.0	24	24	ABQ02995	Oligonucleotide ad
c	8	13.4	67.0	24	24	ABQ09847	Oligonucleotide ad
	9	13.4	67.0	24	24	ABQ09888	Oligonucleotide ad
c	10	13.4	67.0	41	19	AAV47814	Maize polymorphic
c	11	13.4	67.0	41	19	AAV47815	Maize polymorphic
c	12	13.4	67.0	41	19	AAV47816	Maize polymorphic
c	13	13.4	67.0	41	19	AAV47817	Maize polymorphic
	14	13.4	67.0	88	22	AAS21786	Human collagen gen
	15	13.2	66.0	30	21	AAA74506	Human BAFF cDNA PC
c	16	13.2	66.0	45	12	AAQ11830	Probe to minisatel
c	17	13.2	66.0	45	12	AAQ12265	Probe to minisatel
c	18	13.2	66.0	45	12	AAQ12266	Probe to minisatel
c	19	13.2	66.0	45	12	AAQ12267	Probe to minisatel
c	20	13.2	66.0	45	12	AAQ12268	Probe to minisatel
c	21	13.2	66.0	45	12	AAQ12269	Probe to minisatel
c	22	13.2	66.0	45	12	AAQ12270	Probe to minisatel
c	23	13.2	66.0	45	12	AAQ12271	Probe to minisatel
c	24	13.2	66.0	45	12	AAQ12272	Probe to minisatel
c	25	13.2	66.0	45	12	AAQ12273	Probe to minisatel
c	26	13.2	66.0	45	12	AAQ12274	Probe to minisatel
c	27	13.2	66.0	45	12	AAQ12275	Probe to minisatel
c	28	13.2	66.0	45	12	AAQ12277	Probe to minisatel
c	29	13.2	66.0	45	12	AAQ12279	Probe to minisatel
c	30	13.2	66.0	45	12	AAQ12280	Probe to minisatel
c	31	13.2	66.0	45	12	AAQ12281	Probe to minisatel
c	32	13.2	66.0	45	12	AAQ12282	Probe to minisatel
c	33	13.2	66.0	45	12	AAQ12283	Probe to minisatel
c	34	13.2	66.0	51	18	AAV76578	Staphylococcus aur
	35	13.2	66.0	60	24	ABN41114	Human spliced tran
c	36	13.2	66.0	60	24	ABN45600	Human spliced tran
c	37	13.2	66.0	65	24	ABN53857	Mouse spliced tran
	38	13.2	66.0	90	24	ABK36765	Human DNA encoding
c	39	13.2	66.0	98	16	AAQ84550	Mycobacterium simi
	40	13.2	66.0	99	22	ABA72855	Human foetal liver
	41	13.2	66.0	99	22	AAK21288	Human brain expres
	42	13.2	66.0	99	22	AAK47444	Human bone marrow
	43	13	65.0	20	21	AAA59802	Primer for CSF1 re
	44	13	65.0	20	22	AAH41784	CSF1 receptor-2 ge
	45	13	65.0	49	17	AAT35845	Primer for human i

ALIGNMENTS

RESULT 1
AAZ31444
ID AAZ31444 standard; DNA; 20 BP.
XX
AC AAZ31444;
XX
DT 07-FEB-2000 (first entry)
XX
DE Human neuropilin mRNA specific antisense oligo GTI3614.
XX
KW Neuropilin; human; growth; metastasis; tumor; neovascularisation;
KW cancer; papilloma; diabetic retinopathy; antisense; ss.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9955855-A2.
XX
PD 04-NOV-1999.
XX
PF 23-APR-1999; 99WO-CA00324.
XX
PR 23-APR-1998; 98US-0082791.
XX
PA (GENE-) GENESENSE TECHNOLOGIES INC.
XX
PI Wright JA, Young AH, Lee YS;
XX
DR WPI; 2000-023357/02.
XX
PT Antisense oligonucleotides that inhibit neuropilin expression, useful
PT for treating cancer -

```
XX
PS   Claim 4; Page 16; 57pp; English.
XX
CC   Sequences AAZ31431-460 represent antisense oligonucleotides which
CC   inhibit human neuropilin expression. The antisense oligonucleotides can
CC   be used to inhibit the growth or metastasis of a mammalian tumor and
CC   inhibit neovascularisation. The oligonucleotides may be used to treat
CC   various forms of cancers or tumors, such as sarcomas, melanomas,
CC   adenomas, carcinomas of solid tissue, hypoxic tumors, squamous cell
CC   carcinomas of the mouth, throat, larynx and lung, genitourinary cancers
CC   such as cervical and bladder cancer, hematopoietic cancers, colon cancer,
CC   breast cancer, pancreatic cancer, renal cancer, brain cancer, skin
CC   cancer, liver cancer, head and neck cancers, and nervous system cancers,
CC   as well as benign lesions such as papillomas. The methods may be used to
CC   treat neovascularisation disorders such as diabetic retinopathy, and
CC   retinopathy of prematurity and age related macular degeneration.
XX
SQ   Sequence 20 BP; 4 A; 7 C; 3 G; 6 T; 0 other;

  Query Match             100.0%;  Score 20;  DB 21;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 1.1;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 TGCTTCCCACCCTGAATGAT 20
           ||||||||||||||||||||
Db       1 TGCTTCCCACCCTGAATGAT 20


RESULT 2
ABN43139/c
ID   ABN43139 standard; DNA; 60 BP.
XX
AC   ABN43139;
XX
DT   15-JUL-2002  (first entry)
XX
DE   Human spliced transcript detection oligonucleotide SEQ ID NO:15887.
XX
KW   Human; mouse; rat; splice transcript; detection; RNA transcript;
KW   splice variant; transcriptome; oligonucleotide library; ss.
XX
OS   Homo sapiens.
XX
PN   WO200210449-A2.
XX
PD   07-FEB-2002.
XX
PF   20-JUL-2001; 2001WO-IB01903.
XX
PR   28-JUL-2000; 2000US-221607P.
PR   02-MAY-2001; 2001US-287724P.
XX
PA   (COMP-) COMPUGEN INC.
XX
PI   Shoshan A,  Wasserman A,  Mintz E,  Mintz L,  Faigler S;
XX
DR   WPI; 2002-257383/30.
XX
PT   New oligonucleotide libraries comprising oligonucleotides which
PT   selectively hybridize to mRNAs transcribed from a transcription unit of
PT   a genome, useful for detecting tissue-, pathology-, and
PT   developmental-specific genes -
XX
PS   Example 1; SEQ ID 15887; 47pp; English.
XX
CC   The present invention describes oligonucleotide libraries for detecting
CC   messenger RNAs that populate a (sub-)transcriptome, where the
CC   (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC   transcription units that populate a genome. The library comprises
CC   several oligonucleotides, each capable of hybridising selectively to a
CC   set of messenger RNAs transcribed from a given transcription unit of
CC   the genome, which encodes one or more messenger RNA splice variants.
CC   The oligonucleotide libraries are useful for detecting mRNAs from a
CC   biological sample, in expression profiling studies, in qualitatively or
CC   quantitatively characterising the corresponding transcriptome, and in
CC   detecting RNA transcripts and splice variants of human or animal
CC   transcriptomes. The libraries may also be used as specialised mini
CC   libraries to detect transcripts of a sub-transcriptome under a
CC   particular biological or pathological state, and so allowing the
CC   detection of tissue- and pathology-specific genes such as those genes
CC   only expressed in specific tissue under a specific pathological
CC   condition; to detect developmental specific genes; and to detect RNA
CC   transcripts and splice variants of a transcriptome of a patient suffering
CC   from a particular disorder. ABN27253 to ABN59589 represent
CC   oligonucleotide sequences from rats, humans and mice, which are used in
CC   the exemplification of the present invention.
CC   N.B. The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 60 BP; 18 A; 9 C; 22 G; 11 T; 0 other;

  Query Match             100.0%;  Score 20;  DB 24;  Length 60;
  Best Local Similarity   100.0%;  Pred. No. 1.3;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 TGCTTCCCACCCTGAATGAT 20
           ||||||||||||||||||||
Db      31 TGCTTCCCACCCTGAATGAT 12


RESULT 3
AAT99600
ID   AAT99600 standard; DNA; 30 BP.
XX
AC   AAT99600;
XX
DT   06-JUL-1998  (first entry)
XX
DE   Protein kinase LIMK1 gene intron 1/exon 2 junction.
XX
KW   Williams syndrome cognitive profile; WSCP; cognition; LIM-kinase 1;
KW   LIMK1 gene; supra-vascular aortic stenosis; protein kinase; human;
KW   ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   intron          1..20
FT                   /*tag= a
FT                   /note= "3' end of intron 1"
FT   exon            21..30
FT                   /*tag= b
FT                   /note= "5' end of exon 2"
XX
PN   WO9801740-A2.
XX
PD   15-JAN-1998.
XX
PF   07-JUL-1997;   97WO-US11687.
XX
PR   10-JUL-1996;   96US-0678039.
XX
PA   (UTAH ) UNIV UTAH RES FOUND.
XX
PI   Keating MT,  Morris CA;
XX
DR   WPI; 1998-101185/09.
XX
PT   Diagnosing Williams syndrome cognitive profile from hemizygosity of
PT   LIMK1 - gene on chromosome 7 encoding new kinase, allowing
PT   differentiation from classic Williams syndrome and supra-vascular
PT   aortic stenosis
XX
PS   Example 3; Page 24; 62pp; English.
XX
```

CC This sequence comprises the junction between intron 1 and exon 2 of
CC the human LIM-kinase 1 (LIMK1) gene. Exon 2 contains 97 coding
CC nucleotides. The LIMK1 genomic structure was determined by DNA
CC sequence analyses of genomic clones containing LIMK1. The gene
CC consists of 16 exons (see also AAV05315 and AAT99599-T99629) and spans
CC approximately 37 kb. The gene is located 15.4 kb 3' of the
CC elastin gene in chromosome 7. It encodes a novel protein kinase
CC (see AAW46576). A claimed method for determining the presence of
CC Williams syndrome cognitive profile (WSCP) comprises determining
CC the zygosity of the LIMK1 locus, with hemizygosity being indicative
CC of impaired visuo-spatial constructive cognition. A claimed method
CC for distinguishing whether an individual has Williams syndrome
CC (WS), WSCP or SVAS (supra-vascular aortic stenosis) involves
CC analysis of deletion of parts of chromosome 7. Deletion of the ELN
CC (elastin) locus but not LIMK1 indicates SVAS, deletion of ELN and
CC LIMK1 but no more than about 100 kb 3' to LIMK1 indicates WSCP, and
CC deletion of ELN, LIMK1 and over 300 kb 3' of LIMK1 indicates WS.
XX
SQ Sequence 30 BP; 6 A; 12 C; 6 G; 6 T; 0 other;

Query Match 69.0%; Score 13.8; DB 19; Length 30;
Best Local Similarity 88.2%; Pred. No. 1.6e+03;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 CTTCCCACCCTGAATGA 19
||||||||||||| | ||
Db 6 CTTCCCACCCTGCAGGA 22

RESULT 4
AAT44095
ID AAT44095 standard; DNA; 34 BP.
XX
AC AAT44095;
XX
DT 28-FEB-1997 (first entry)
XX
DE Human G-protein thrombin-like receptor 5' primer.
XX
KW G-protein; thrombin; receptor; diagnosis; gene therapy;
KW haemophilia; wound healing; restenosis; angina; inflammation;
KW polymerase chain reaction; PCR; primer; COS-7; ss.
XX
OS Synthetic.
XX
PN WO9639438-A1.
XX
PD 12-DEC-1996.
XX
PF 06-JUN-1995; 95WO-US07180.
XX
PR 06-JUN-1995; 95WO-US07180.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Gocayne JD, Li Y, Ruben SM;
XX
DR WPI; 1997-043073/04.
XX
PT Human G-protein thrombin receptor, HIBEB69 - useful to identify
PT (ant)agonists, for treatment of angina, restenosis, wound healing
PT etc.
XX
PS Example 2; Page 36; 58pp; English.
XX
CC A PCR primer (AAT44095) contains a HindIII site followed by 18
CC nucleotides of the human G-protein thrombin-like receptor coding
CC sequence (see also AAT44092) starting from the initiation codon. It
CC was used with a 3' primer (AAT44096) contg. HA tag complementary
CC sequences to amplify a G-protein thrombin-like receptor DNA clone.
CC The PCR product was ligated into vector pcDNAI/Amp to allow prodn.
CC of the recombinant receptor (see also AAW07617) in transfected COS
CC cells, and its purification using HA-specific monoclonal antibody.
XX
SQ Sequence 34 BP; 8 A; 9 C; 9 G; 8 T; 0 other;

Query Match 69.0%; Score 13.8; DB 18; Length 34;
Best Local Similarity 88.2%; Pred. No. 1.6e+03;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 GCTTCCCACCCTGAATG 18
|||| ||||| ||||||
Db 7 GCTTGCCACCATGAATG 23

RESULT 5
AAQ11829/c
ID AAQ11829 standard; DNA; 45 BP.
XX
AC AAQ11829;
XX
DT 30-JUL-1991 (first entry)
XX
DE Probe to minisatellite sequence upstream of human alpha 2 plasmin
DE inhibitor gene.
XX
KW Genetic fingerprinting; ss.
XX
OS Synthetic.
XX
PN WO9106675-A.
XX
PD 16-MAY-1991.
XX
PF 31-OCT-1990; 90WO-JP01404.
XX
PR 02-NOV-1989; 89JP-0284909.
XX
PA (FARH) HOECHST JAPAN LTD.
XX
PI Hashimoto T, Kurose K;
XX
DR WPI; 1991-164221/22.
XX
PT 45-base oligo:nucleotide probe for DNA fingerprinting - detects
PT mini-satellite repeat sequence upstream of the alpha-2-plasmin
PT inhibitor gene.
XX
PS Claim 2; Page 13; 20pp; Japanese.
XX
CC The minisatelite sequence is repeated from 12 to 20 times, the number
CC being stably inherited. Probes to the sequence may be used in
CC Southern blot analyses for kinship tests eg. DNA fingerprinting.
XX
SQ Sequence 45 BP; 8 A; 4 C; 21 G; 5 T; 7 other;

Query Match 68.0%; Score 13.6; DB 12; Length 45;
Best Local Similarity 72.2%; Pred. No. 2.1e+03;
Matches 13; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 TGCTTCCCACCCTGAATG 18
::|:|||||||||| | ||
Db 23 YSCYTCCCACCCTCACTG 6

RESULT 6
ABL00525/c
ID ABL00525 standard; DNA; 51 BP.
XX
AC ABL00525;
XX
DT 05-MAR-2002 (first entry)
XX
DE Human silent noncoding SNP oligonucleotide SEQ ID NO:516.
XX
KW Human; single nucleotide polymorphism; SNP; polymorphism; cytostatic;

KW immunosuppressive; antiinflammatory; neuroprotective; antimicrobial;
KW autoimmune disease; inflammation; cancer; nervous system disease;
KW infection; polymorphic protein; ds.
XX
OS Homo sapiens.
XX
PN WO200138586-A2.
XX
PD 31-MAY-2001.
XX
PF 22-NOV-2000; 2000WO-US32311.
XX
PR 24-NOV-1999; 99US-0167383.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shimkets RA, Leach M;
XX
DR WPI; 2001-355949/37.
XX
PT Isolated human nucleic acids comprising one or more single nucleotide
PT polymorphisms, useful for treating a subject suffering from a
PT pathology, e.g. autoimmune diseases, ascribed to the presence of a
PT sequence polymorphism -
XX
PS Claim 1; Page 403; 674pp; English.
XX
CC ABL00010 to ABL01104 represent human nucleic acid oligonucleotides
CC comprising one or more single nucleotide polymorphisms (SNPs). ABB56531
CC to ABB56903 represent human peptides encoded by some of the SNP
CC oligonucleotides. The sequences from the present invention can have
CC immunosuppressive, cytostatic, antiinflammatory, neuroprotective and
CC antimicrobial activities. Nucleic acids, polypeptides, oligonucleotides
CC and antibodies from the present invention can be used for treating a
CC subject suffering from, at risk for, or suspected of, suffering from a
CC pathology ascribed to the presence of a sequence polymorphism. The
CC pathology may be autoimmune diseases, inflammation, cancer, diseases of
CC the nervous system, and infection by pathogenic microorganisms. The SNPs
CC are also useful for determining which forms of a characterised
CC polymorphism are present in individuals. The antibodies may be used in
CC the detection, quantitation and/or cellular or tissue localisation of a
CC polymorphic protein (e.g., for use in measuring levels of the
CC polymorphic protein within appropriate physiological samples).
XX
SQ Sequence 51 BP; 11 A; 12 C; 11 G; 17 T; 0 other;

Query Match 68.0%; Score 13.6; DB 23; Length 51;
Best Local Similarity 80.0%; Pred. No. 2.2e+03;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
Qy       1 TGCTTCCCACCCTGAATGAT 20
           || | | ||||||| |||||
Db      36 TGTTCCACACCCTGGATGAT 17
```

RESULT 7
ABQ02995/c
ID ABQ02995 standard; DNA; 24 BP.
XX
AC ABQ02995;
XX
DT 11-JUN-2002 (first entry)
XX
DE Oligonucleotide adapter/capture probe 2986.
XX
KW Oligonucleotide array; adapter sequence; probe; ss.
XX
OS Synthetic.
XX
PN WO200216649-A2.
XX
PD 28-FEB-2002.
XX
PF 27-AUG-2001; 2001WO-US26519.
XX
PR 25-AUG-2000; 2000US-227948P.
PR 29-AUG-2000; 2000US-228854P.
XX
PA (ILLU-) ILLUMINA INC.
XX
PI Gunderson K;
XX
DR WPI; 2002-292068/33.
XX
PT Array comprising adapter sequences useful for immobilizing or detecting
PT a target nucleic acid sequence, has different addresses comprising
PT different specific capture probes -
XX
PS Claim 1; Page 115; 261pp; English.
XX
CC The invention relates to an oligonucleotide array (I) comprising at least
CC 25 different addresses (adapter sequences) with each comprising a
CC different capture probe selected from a group consisting of the sequences
CC given in ABQ00010-ABQ13409. (I) is useful for immobilising a target
CC nucleic acid sequence by attaching a adapter nucleic acid
CC (ABQ00010-ABQ13409) to a target nucleic acid to form a modified target
CC nucleic acid and contacting the modified target nucleic acid with (I).
CC The steps of above method is useful for detecting a target nucleic acid,
CC which further comprises detecting the presence of the modified target
CC nucleic acid.
XX
SQ Sequence 24 BP; 5 A; 6 C; 8 G; 5 T; 0 other;

Query Match 67.0%; Score 13.4; DB 24; Length 24;
Best Local Similarity 93.3%; Pred. No. 2.5e+03;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy       5 TCCCACCCTGAATGA 19
           ||||||||||| |||
Db      21 TCCCACCCTGATTGA 7
```

RESULT 8
ABQ09847/c
ID ABQ09847 standard; DNA; 24 BP.
XX
AC ABQ09847;
XX
DT 11-JUN-2002 (first entry)
XX
DE Oligonucleotide adapter/capture probe 9838.
XX
KW Oligonucleotide array; adapter sequence; probe; ss.
XX
OS Synthetic.
XX
PN WO200216649-A2.
XX
PD 28-FEB-2002.
XX
PF 27-AUG-2001; 2001WO-US26519.
XX
PR 25-AUG-2000; 2000US-227948P.
PR 29-AUG-2000; 2000US-228854P.
XX
PA (ILLU-) ILLUMINA INC.
XX
PI Gunderson K;
XX
DR WPI; 2002-292068/33.
XX
PT Array comprising adapter sequences useful for immobilizing or detecting
PT a target nucleic acid sequence, has different addresses comprising
PT different specific capture probes -
XX
PS Claim 1; Page 210; 261pp; English.

```
XX
CC   The invention relates to an oligonucleotide array (I) comprising at least
CC   25 different addresses (adapter sequences) with each comprising a
CC   different capture probe selected from a group consisting of the sequences
CC   given in ABQ00010-ABQ13409. (I) is useful for immobilising a target
CC   nucleic acid sequence by attaching a adapter nucleic acid
CC   (ABQ00010-ABQ13409) to a target nucleic acid to form a modified target
CC   nucleic acid and contacting the modified target nucleic acid with (I).
CC   The steps of above method is useful for detecting a target nucleic acid,
CC   which further comprises detecting the presence of the modified target
CC   nucleic acid.
XX
SQ   Sequence 24 BP; 5 A; 6 C; 8 G; 5 T; 0 other;

  Query Match             67.0%;  Score 13.4;  DB 24;  Length 24;
  Best Local Similarity   93.3%;  Pred. No. 2.5e+03;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       5 TCCCACCCTGAATGA 19
           ||||||||||| |||
Db      21 TCCCACCCTGATTGA 7


RESULT 9
ABQ09888
ID   ABQ09888 standard; DNA; 24 BP.
XX
AC   ABQ09888;
XX
DT   11-JUN-2002  (first entry)
XX
DE   Oligonucleotide adapter/capture probe 9879.
XX
KW   Oligonucleotide array; adapter sequence; probe; ss.
XX
OS   Synthetic.
XX
PN   WO200216649-A2.
XX
PD   28-FEB-2002.
XX
PF   27-AUG-2001; 2001WO-US26519.
XX
PR   25-AUG-2000; 2000US-227948P.
PR   29-AUG-2000; 2000US-228854P.
XX
PA   (ILLU-) ILLUMINA INC.
XX
PI   Gunderson K;
XX
DR   WPI; 2002-292068/33.
XX
PT   Array comprising adapter sequences useful for immobilizing or detecting
PT   a target nucleic acid sequence, has different addresses comprising
PT   different specific capture probes -
XX
PS   Claim 1; Page 210; 261pp; English.
XX
CC   The invention relates to an oligonucleotide array (I) comprising at least
CC   25 different addresses (adapter sequences) with each comprising a
CC   different capture probe selected from a group consisting of the sequences
CC   given in ABQ00010-ABQ13409. (I) is useful for immobilising a target
CC   nucleic acid sequence by attaching a adapter nucleic acid
CC   (ABQ00010-ABQ13409) to a target nucleic acid to form a modified target
CC   nucleic acid and contacting the modified target nucleic acid with (I).
CC   The steps of above method is useful for detecting a target nucleic acid,
CC   which further comprises detecting the presence of the modified target
CC   nucleic acid.
XX
SQ   Sequence 24 BP; 5 A; 8 C; 6 G; 5 T; 0 other;

  Query Match             67.0%;  Score 13.4;  DB 24;  Length 24;
  Best Local Similarity   93.3%;  Pred. No. 2.5e+03;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       5 TCCCACCCTGAATGA 19
           ||||||||||| |||
Db       4 TCCCACCCTGATTGA 18


RESULT 10
AAV47814/c
ID   AAV47814 standard; DNA; 41 BP.
XX
AC   AAV47814;
XX
DT   14-OCT-1998  (first entry)
XX
DE   Maize polymorphic site oligonucleotide marker UMC109-G2/G3-1A.
XX
KW   Maize; marker; probe; PCR primer; polymorphism; vegetal sequence;
KW   polymorphic site; corn; graminae species; ss.
XX
OS   Synthetic.
OS   Zea sp.
XX
PN   WO9830717-A2.
XX
PD   16-JUL-1998.
XX
PF   02-DEC-1997;   97WO-EP07134.
XX
PR   02-DEC-1996;   96US-0032069.
XX
PA   (BIOC-) BIOCEM SA.
XX
PI   Murigneux A;
XX
DR   WPI; 1998-399160/34.
XX
PT   Vegetal sequences including single nucleotide polymorphism - useful,
PT   e.g. to determine polymorphisms in plants, determine strain in plant
PT   breeding and to correlate polymorphisms with phenotypic traits
XX
PS   Claim 2; Page 13; 32pp; English.
XX
CC   The present invention describes a nucleic acid segment comprising at
CC   least 10 contiguous nucleotides from a vegetal sequence including a
CC   polymorphic site which is a single nucleotide polymorphism (SNP), or the
CC   complement of the segment. Also described are: (1) an allele-specific
CC   oligonucleotides hybridising to segment, or their complements, and (2) a
CC   method of analysing nucleic acids from a subject, by determining if a
CC   base is occupying any one (or a set) of polymorphic sites in 261
CC   sequences derived from six maize lines (see AAV47701 to AAV47961). The
CC   segments are useful in fingerprint analysis in plants to determine which
CC   polymorphisms are present, which strain a plant belongs to and to
CC   distinguish between strains. The polymorphisms may correlate with
CC   phenotypic traits (e.g. plant growth rate or crop yield), and the
CC   segments are useful to determine the presence/absence of specific
CC   polymorphisms correlating with the existence/absence of particular
CC   traits. The segments are also useful in marker assisted back-cross
CC   techniques to select plants with a higher percentage of recurrent parent
CC   in a back-cross population. Segments incorporate SNPs which occur more
CC   frequently than other polymorphism types and are therefore more likely
CC   to be located close to genetic loci of interest; different forms of
CC   characterised SNPs are also often easier to detect than other
CC   polymorphism types.
XX
SQ   Sequence 41 BP; 7 A; 10 C; 16 G; 7 T; 1 other;

  Query Match             67.0%;  Score 13.4;  DB 19;  Length 41;
  Best Local Similarity   82.4%;  Pred. No. 2.6e+03;
  Matches   14;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 TGCTTCCCACCCTGAAT 17
           ||||||: ||||| |||
```

```
Db     27 TGCTTCSGACCCTCAAT 11


RESULT 11
AAV47815/c
ID   AAV47815 standard; DNA; 41 BP.
XX
AC   AAV47815;
XX
DT   14-OCT-1998  (first entry)
XX
DE   Maize polymorphic site oligonucleotide marker UMC109-G2/G3-1B.
XX
KW   Maize; marker; probe; PCR primer; polymorphism; vegetal sequence;
KW   polymorphic site; corn; graminae species; ss.
XX
OS   Synthetic.
OS   Zea sp.
XX
PN   WO9830717-A2.
XX
PD   16-JUL-1998.
XX
PF   02-DEC-1997;   97WO-EP07134.
XX
PR   02-DEC-1996;   96US-0032069.
XX
PA   (BIOC-) BIOCEM SA.
XX
PI   Murigneux A;
XX
DR   WPI; 1998-399160/34.
XX
PT   Vegetal sequences including single nucleotide polymorphism - useful,
PT   e.g. to determine polymorphisms in plants, determine strain in plant
PT   breeding and to correlate polymorphisms with phenotypic traits
XX
PS   Claim 2; Page 13; 32pp; English.
XX
CC   The present invention describes a nucleic acid segment comprising at
CC   least 10 contiguous nucleotides from a vegetal sequence including a
CC   polymorphic site which is a single nucleotide polymorphism (SNP), or the
CC   complement of the segment. Also described are: (1) an allele-specific
CC   oligonucleotides hybridising to segment, or their complements, and (2) a
CC   method of analysing nucleic acids from a subject, by determining if a
CC   base is occupying any one (or a set) of polymorphic sites in 261
CC   sequences derived from six maize lines (see AAV47701 to AAV47961). The
CC   segments are useful in fingerprint analysis in plants to determine which
CC   polymorphisms are present, which strain a plant belongs to and to
CC   distinguish between strains. The polymorphisms may correlate with
CC   phenotypic traits (e.g. plant growth rate or crop yield), and the
CC   segments are useful to determine the presence/absence of specific
CC   polymorphisms correlating with the existence/absence of particular
CC   traits. The segments are also useful in marker assisted back-cross
CC   techniques to select plants with a higher percentage of recurrent parent
CC   in a back-cross population. Segments incorporate SNPs which occur more
CC   frequently than other polymorphism types and are therefore more likely
CC   to be located close to genetic loci of interest; different forms of
CC   characterised SNPs are also often easier to detect than other
CC   polymorphism types.
XX
SQ   Sequence 41 BP; 7 A; 10 C; 15 G; 8 T; 1 other;

  Query Match             67.0%;  Score 13.4;  DB 19;  Length 41;
  Best Local Similarity   82.4%;  Pred. No. 2.6e+03;
  Matches   14;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 TGCTTCCCACCCTGAAT 17
          ||||||: ||||| |||
Db     27 TGCTTCSGACCCTCAAT 11


RESULT 12
```

```
AAV47816/c
ID   AAV47816 standard; DNA; 41 BP.
XX
AC   AAV47816;
XX
DT   14-OCT-1998  (first entry)
XX
DE   Maize polymorphic site oligonucleotide marker UMC109-G2/G3-1C.
XX
KW   Maize; marker; probe; PCR primer; polymorphism; vegetal sequence;
KW   polymorphic site; corn; graminae species; ss.
XX
OS   Synthetic.
OS   Zea sp.
XX
PN   WO9830717-A2.
XX
PD   16-JUL-1998.
XX
PF   02-DEC-1997;   97WO-EP07134.
XX
PR   02-DEC-1996;   96US-0032069.
XX
PA   (BIOC-) BIOCEM SA.
XX
PI   Murigneux A;
XX
DR   WPI; 1998-399160/34.
XX
PT   Vegetal sequences including single nucleotide polymorphism - useful,
PT   e.g. to determine polymorphisms in plants, determine strain in plant
PT   breeding and to correlate polymorphisms with phenotypic traits
XX
PS   Claim 2; Page 13; 32pp; English.
XX
CC   The present invention describes a nucleic acid segment comprising at
CC   least 10 contiguous nucleotides from a vegetal sequence including a
CC   polymorphic site which is a single nucleotide polymorphism (SNP), or the
CC   complement of the segment. Also described are: (1) an allele-specific
CC   oligonucleotides hybridising to segment, or their complements, and (2) a
CC   method of analysing nucleic acids from a subject, by determining if a
CC   base is occupying any one (or a set) of polymorphic sites in 261
CC   sequences derived from six maize lines (see AAV47701 to AAV47961). The
CC   segments are useful in fingerprint analysis in plants to determine which
CC   polymorphisms are present, which strain a plant belongs to and to
CC   distinguish between strains. The polymorphisms may correlate with
CC   phenotypic traits (e.g. plant growth rate or crop yield), and the
CC   segments are useful to determine the presence/absence of specific
CC   polymorphisms correlating with the existence/absence of particular
CC   traits. The segments are also useful in marker assisted back-cross
CC   techniques to select plants with a higher percentage of recurrent parent
CC   in a back-cross population. Segments incorporate SNPs which occur more
CC   frequently than other polymorphism types and are therefore more likely
CC   to be located close to genetic loci of interest; different forms of
CC   characterised SNPs are also often easier to detect than other
CC   polymorphism types.
XX
SQ   Sequence 41 BP; 7 A; 10 C; 15 G; 8 T; 1 other;

  Query Match             67.0%;  Score 13.4;  DB 19;  Length 41;
  Best Local Similarity   82.4%;  Pred. No. 2.6e+03;
  Matches   14;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 TGCTTCCCACCCTGAAT 17
          ||||||: ||||| |||
Db     27 TGCTTCSGACCCTCAAT 11


RESULT 13
AAV47817/c
ID   AAV47817 standard; DNA; 41 BP.
XX
AC   AAV47817;
```

XX
DT 14-OCT-1998 (first entry)
XX
DE Maize polymorphic site oligonucleotide marker UMC109-G2/G3-1D.
XX
KW Maize; marker; probe; PCR primer; polymorphism; vegetal sequence;
KW polymorphic site; corn; graminae species; ss.
XX
OS Synthetic.
OS Zea sp.
XX
PN WO9830717-A2.
XX
PD 16-JUL-1998.
XX
PF 02-DEC-1997; 97WO-EP07134.
XX
PR 02-DEC-1996; 96US-0032069.
XX
PA (BIOC-) BIOCEM SA.
XX
PI Murigneux A;
XX
DR WPI; 1998-399160/34.
XX
PT Vegetal sequences including single nucleotide polymorphism - useful,
PT e.g. to determine polymorphisms in plants, determine strain in plant
PT breeding and to correlate polymorphisms with phenotypic traits
XX
PS Claim 2; Page 13; 32pp; English.
XX
CC The present invention describes a nucleic acid segment comprising at
CC least 10 contiguous nucleotides from a vegetal sequence including a
CC polymorphic site which is a single nucleotide polymorphism (SNP), or the
CC complement of the segment. Also described are: (1) an allele-specific
CC oligonucleotides hybridising to segment, or their complements, and (2) a
CC method of analysing nucleic acids from a subject, by determining if a
CC base is occupying any one (or a set) of polymorphic sites in 261
CC sequences derived from six maize lines (see AAV47701 to AAV47961). The
CC segments are useful in fingerprint analysis in plants to determine which
CC polymorphisms are present, which strain a plant belongs to and to
CC distinguish between strains. The polymorphisms may correlate with
CC phenotypic traits (e.g. plant growth rate or crop yield), and the
CC segments are useful to determine the presence/absence of specific
CC polymorphisms correlating with the existence/absence of particular
CC traits. The segments are also useful in marker assisted back-cross
CC techniques to select plants with a higher percentage of recurrent parent
CC in a back-cross population. Segments incorporate SNPs which occur more
CC frequently than other polymorphism types and are therefore more likely
CC to be located close to genetic loci of interest; different forms of
CC characterised SNPs are also often easier to detect than other
CC polymorphism types.
XX
SQ Sequence 41 BP; 7 A; 10 C; 14 G; 9 T; 1 other;

Query Match 67.0%; Score 13.4; DB 19; Length 41;
Best Local Similarity 82.4%; Pred. No. 2.6e+03;
Matches 14; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 TGCTTCCCACCCTGAAT 17
|||||| ||||| |||
Db 27 TGCTTCSGACCCTCAAT 11

RESULT 14
AAS21786
ID AAS21786 standard; DNA; 88 BP.
XX
AC AAS21786;
XX
DT 24-OCT-2001 (first entry)
XX
DE Human collagen gene COL1A1intron 12.
XX
KW Human; collagen; COL1A1; COL1A2; COL9A1; COL9A2; COL9A3; ds;
KW osteoporosis; multiple epiphyseal dysplasia; osteogenesis imperfecta;
KW shortness of stature; low bone density; gene therapy; intron.
XX
OS Homo sapiens.
XX
PN US6265157-B1.
XX
PD 24-JUL-2001.
XX
PF 03-OCT-1997; 97US-0943731.
XX
PR 03-DEC-1991; 91US-0803628.
PR 13-MAR-1994; 94US-0212322.
XX
PA (UYAL-) UNIV ALLEGHENY HEALTH SCI.
PA (UYJE-) UNIV JEFFERSON THOMAS.
PA (UYOU-) UNIV OULU.
XX
PI Prockop DJ, Spotila LD, Deltas CD, Sereda L, Westerhausen Larson A;
PI Pack M, Colige A, Early J, Koerkkoe J, Ala-Kokko L, Annunen S;
PI Pihlajamaa T, Vuoristo M, Paassilta P;
XX
DR WPI; 2001-432201/46.
XX
PT Detecting collagen gene alteration, useful for diagnosing osteoporosis,
PT multiple epiphyseal dysplasia, osteogenesis imperfecta, shortness of
PT stature and low bone density in humans -
XX
PS Example 4; Fig 4B; 617pp; English.
XX
CC The invention relates to Detecting a collagen gene alteration associated
CC with a pathological condition in a human subject by obtaining from the
CC subject a sample nucleic acid containing a portion of at least 15
CC consecutive nucleotides of the segment of the COL1A1 gene extending in
CC the 5' to 3' direction from 78 nucleotides of intron 27 located adjacent
CC exon 28 through the 3' end of intron 51, where the portion contains an
CC intronic nucleotide and a first and second site, determining the sequence
CC of the portion and comparing the sequence of the portion with the
CC corresponding consensus sequence of the COL1A1 gene where a difference
CC between the sequence of the portion and the consensus sequence indicates
CC the presence of the collagen alteration in the subject. The method is
CC used for detecting abnormalities in a COL1 or COL9 gene is useful for
CC determining whether a subject is afflicted with pathological conditions
CC associated with an altered collagen gene such as osteoporosis, multiple
CC epiphyseal dysplasia, osteogenesis imperfecta, shortness of stature and
CC low bone density. Identification of an abnormality in a collagen gene is
CC also useful for designing a therapeutic nucleotide or gene therapy agent
CC which can be administered to the subject to correct or alleviate the
CC abnormality. The method is useful for detecting mutations in both the
CC coding and non-coding sequences of any of the COL1 or COL9 genes.
CC Therefore the method can be used to detect collagen gene alterations
CC which affect either the primary sequence of a collagen protein chain,
CC splicing of the mRNA encoding such chains or regulation of expression of
CC the genes encoding such chains. The present sequence is an intron from a
CC collagen gene of the invention.
XX
SQ Sequence 88 BP; 23 A; 26 C; 14 G; 25 T; 0 other;

Query Match 67.0%; Score 13.4; DB 22; Length 88;
Best Local Similarity 93.3%; Pred. No. 2.9e+03;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 5 TCCCACCCTGAATGA 19
||||||| |||||||
Db 44 TCCCACCATGAATGA 58

RESULT 15
AAA74506
ID AAA74506 standard; DNA; 30 BP.
XX

AC AAA74506;
XX
DT 12-DEC-2000 (first entry)
XX
DE Human BAFF cDNA PCR primer JT1069.
XX
KW Human; BAFF; PCR primer; B-cell co-stimulation; B-cell growth;
KW B cell activating factor belonging to the TNF family;
KW immunoglobulin secretion; autoimmune disease; tumour; hypertension;
KW inflammation; immunosuppressive disease; HIV; organ transplantation; ss.
XX
OS Homo sapiens.
XX
PN WO200043032-A2.
XX
PD 27-JUL-2000.
XX
PF 25-JAN-2000; 2000WO-US01788.
XX
PR 25-JAN-1999; 99US-0117169.
PR 09-JUL-1999; 99US-0143228.
XX
PA (BIOJ) BIOGEN INC.
PA (APOT-) APOTECH SA.
XX
PI Browning J, Ambrose C, Mackay F, Tschopp J, Schneider P;
XX
DR WPI; 2000-482894/42.
XX
PT Stimulating B-cell growth, immunoglobulin production or dendritic
PT cell-induced B-cell growth and maturation, to treat autoimmune and
PT immunosuppressive disorders -
XX
PS Disclosure; Page 21; 75pp; English.
XX
CC The present sequence is a PCR primer for the coding sequence of human
CC "B cell activating factor belonging to the TNF family" (BAFF). This
CC primer was used to isolate the full-length human BAFF coding sequence.
CC BAFF is a ligand belonging to the TNF cytokine family, and is thought to
CC be expressed by T cells and dendritic cells for B-cell co-stimulation.
CC BAFF may be used to stimulate the growth of B-cells and immunoglobulin
CC secretion. BAFF may be used to treat autoimmune diseases, tumours,
CC hypertension, disorders related to B-cell proliferation and maturation,
CC BAFF ligand regulation and inflammation. Also, BAFF may be used to treat
CC an immunosuppressive disease, e.g. human immunodeficiency virus (HIV)
CC infection and immunosuppression related to organ transplantation.
XX
SQ Sequence 30 BP; 9 A; 10 C; 6 G; 5 T; 0 other;

Query Match 66.0%; Score 13.2; DB 21; Length 30;
Best Local Similarity 83.3%; Pred. No. 3.2e+03;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy      2 GCTTCCCACCCTGAATGA 19
          |||| ||||| || ||||
Db      6 GCTTGCCACCATGGATGA 23
```

Search completed: November 23, 2002, 06:29:12
Job time : 100.6 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 05:53:51 ; Search time 21.55 Seconds
(without alignments)
284.619 Million cell updates/sec

Title: US-09-296-264-14
Perfect score: 20
Sequence: 1 tgcttcccaccctgaatgat 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 687286

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	13.8	69.0	30	2	US-08-678-039A-7	Sequence 7, Appli
	2	13.8	69.0	34	5	PCT-US95-07180-6	Sequence 6, Appli
	3	13.4	67.0	88	4	US-08-943-731-18	Sequence 18, Appl
c	4	13.2	66.0	98	1	US-08-403-762A-172	Sequence 172, App
c	5	13	65.0	19	4	US-09-496-632C-8	Sequence 8, Appli
	6	13	65.0	49	1	US-08-481-003-21	Sequence 21, Appl
	7	13	65.0	49	3	US-08-485-598-21	Sequence 21, Appl
	8	12.8	64.0	21	1	US-07-808-455A-4	Sequence 4, Appli
c	9	12.8	64.0	21	2	US-08-403-888A-71	Sequence 71, Appl
	10	12.8	64.0	21	4	US-08-430-225A-11	Sequence 11, Appl
c	11	12.8	64.0	21	4	US-08-430-225A-14	Sequence 14, Appl
	12	12.8	64.0	35	2	US-08-790-963-43	Sequence 43, Appl
	13	12.8	64.0	35	4	US-09-371-774-43	Sequence 43, Appl
	14	12.8	64.0	95	1	US-08-518-878B-3	Sequence 3, Appli
	15	12.8	64.0	95	1	US-08-294-522B-3	Sequence 3, Appli
	16	12.8	64.0	95	2	US-08-807-861A-3	Sequence 3, Appli
	17	12.8	64.0	95	2	US-08-470-868A-3	Sequence 3, Appli
	18	12.8	64.0	95	3	US-09-210-681-3	Sequence 3, Appli
	19	12.8	64.0	95	3	US-08-946-719A-3	Sequence 3, Appli
	20	12.6	63.0	34	3	US-08-466-343D-5	Sequence 5, Appli
	21	12.6	63.0	35	2	US-08-458-970A-5	Sequence 5, Appli
c	22	12.6	63.0	47	1	US-08-445-050-11	Sequence 11, Appl
c	23	12.6	63.0	47	1	US-08-204-691-11	Sequence 11, Appl
c	24	12.6	63.0	78	2	US-09-073-032-6	Sequence 6, Appli
	25	12.4	62.0	23	4	US-09-467-082-4	Sequence 4, Appli
c	26	12.4	62.0	24	1	US-08-591-070A-43	Sequence 43, Appl
c	27	12.4	62.0	24	2	US-08-927-855-43	Sequence 43, Appl
	28	12.4	62.0	61	4	US-09-302-812-34	Sequence 34, Appl
	29	12.4	62.0	61	4	US-09-511-477-34	Sequence 34, Appl
	30	12.4	62.0	61	4	US-09-511-507-34	Sequence 34, Appl
c	31	12.2	61.0	30	4	US-09-161-466-3	Sequence 3, Appli
c	32	12.2	61.0	30	4	US-09-198-603C-19	Sequence 19, Appl
	33	12.2	61.0	51	4	US-08-218-369-5	Sequence 5, Appli
c	34	12.2	61.0	51	4	US-08-218-369-14	Sequence 14, Appl
	35	12.2	61.0	51	5	PCT-US95-03742-5	Sequence 5, Appli
c	36	12.2	61.0	51	5	PCT-US95-03742-14	Sequence 14, Appl
	37	12	60.0	20	4	US-09-657-481A-25	Sequence 25, Appl
	38	12	60.0	25	1	US-08-155-746-8	Sequence 8, Appli
	39	12	60.0	25	1	US-08-341-148-10	Sequence 10, Appl
	40	12	60.0	25	5	PCT-US94-00771-8	Sequence 8, Appli
	41	12	60.0	25	5	PCT-US94-14096-10	Sequence 10, Appl
c	42	12	60.0	43	1	US-08-392-678-32	Sequence 32, Appl
c	43	12	60.0	43	1	US-08-457-304A-32	Sequence 32, Appl
c	44	12	60.0	43	1	US-08-456-701A-32	Sequence 32, Appl
c	45	12	60.0	43	4	US-08-684-932A-32	Sequence 32, Appl

ALIGNMENTS

RESULT 1
US-08-678-039A-7
; Sequence 7, Application US/08678039A
; Patent No. 5858662
; GENERAL INFORMATION:
; APPLICANT: Keating, Mark T.
; APPLICANT: Morris, Colleen A.
; TITLE OF INVENTION: Diagnosis of Williams Syndrome and
; TITLE OF INVENTION: Williams Syndrome Cognitive Profile by Analysis of the
; TITLE OF INVENTION: Presence or Absence of a LIM-Kinase Gene
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rothwell, Figg, Ernst & Kurz, P.C.
; STREET: 555 Thirteenth Street, N.W., Suite 701 East
; STREET: Tower
; CITY: Washington
; STATE: DC
; COUNTRY: U.S.A.
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/678,039A
; FILING DATE: 10-JUL-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Saxe, Stephen A.
; REGISTRATION NUMBER: 38,609
; REFERENCE/DOCKET NUMBER: 2323-120A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-624-1589
; TELEFAX: 202-783-6031
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 30 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
US-08-678-039A-7

Query Match 69.0%; Score 13.8; DB 2; Length 30;
Best Local Similarity 88.2%; Pred. No. 1.6e+02;

```
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       3 CTTCCCACCCTGAATGA 19
           ||||||||||||  | ||
Db       6 CTTCCCACCCTGCAGGA 22


RESULT 2
PCT-US95-07180-6
; Sequence 6, Application PC/TUS9507180
;  GENERAL INFORMATION:
;    APPLICANT:  LI, YI
;    APPLICANT:  GOCAYNE, JEANINE D
;    APPLICANT:  RUBEN, STEVEN M
;    TITLE OF INVENTION:  G-PROTEIN RECEPTOR HIBEB69
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  CARELLA, BYRNE, BAIN, GILFILLAN, CECCHI,
;      ADDRESSEE:  STEWART & OLSTEIN
;      STREET:  6 BECKER FARM ROAD
;      CITY:  ROSELAND
;      STATE:  NJ
;      COUNTRY:  US
;      ZIP:  07068
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/07180
;      FILING DATE:  06-JUNE-1995
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  MULLINS, J.G.
;      REGISTRATION NUMBER:  30,073
;      REFERENCE/DOCKET NUMBER:  325800-366
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  201-994-1700
;      TELEFAX:  201-994-1744
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  34 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
PCT-US95-07180-6

  Query Match             69.0%;  Score 13.8;  DB 5;  Length 34;
  Best Local Similarity   88.2%;  Pred. No. 1.7e+02;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       2 GCTTCCCACCCTGAATG 18
           |||| ||||| ||||||
Db       7 GCTTGCCACCATGAATG 23


RESULT 3
US-08-943-731-18
; Sequence 18, Application US/08943731
; Patent No. 6265157
;  GENERAL INFORMATION:
;    APPLICANT:  PROCKOP, DARWIN J.
;    APPLICANT:  SPOTILA, LORETTA D.
;    APPLICANT:  DELTAS, CONSTANTINOS D.
;    APPLICANT:  SEREDA, LARISA
;    APPLICANT:  LARSON, ANDREA W.
;    APPLICANT:  PACK, MICHAEL
;    APPLICANT:  COLIGE, ALAIN
;    APPLICANT:  EARLY, JAMES
;    APPLICANT:  KORKKO, JARMO
;    APPLICANT:  ALA-KOKKO, LEENA, et al.
;    TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR DETECTING
;    TITLE OF INVENTION:  ALTERED TYPE I OR TYPE IX COLLAGEN GENE SEQUENCES
;    NUMBER OF SEQUENCES:  666
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  PANITCH SCHWARZE JACOBS & NADEL, P.C.
;      STREET:  ONE COMMERCE SQUARE, 2005 MARKET STREET, 22ND
;      STREET:  FLR.
;      CITY:  PHILADELPHIA
;      STATE:  PA
;      COUNTRY:  USA
;      ZIP:  19103-7086
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/943,731
;      FILING DATE:  03-OCT-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/212,322
;      FILING DATE:  14-MAR-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/803,628
;      FILING DATE:  03-DEC-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DOYLE LEARY Ph.D., KATHRYN
;      REGISTRATION NUMBER:  36,317
;      REFERENCE/DOCKET NUMBER:  9598-27
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  215-965-1284
;      TELEFAX:  215-567-2991
;      TELEX:  831-494
;  INFORMATION FOR SEQ ID NO:  18:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  88 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-943-731-18

  Query Match             67.0%;  Score 13.4;  DB 4;  Length 88;
  Best Local Similarity   93.3%;  Pred. No. 3.1e+02;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       5 TCCCACCCTGAATGA 19
           ||||||| |||||||
Db      44 TCCCACCATGAATGA 58


RESULT 4
US-08-403-762A-172/c
; Sequence 172, Application US/08403762A
; Patent No. 5703217
;  GENERAL INFORMATION:
;    APPLICANT:  MABILAT, Claude
;    APPLICANT:  CHRISTEN, Richard
;    TITLE OF INVENTION:  NUCLEOTIDE FRAGMENT OF THE 23S RIBOSOMAL
;    TITLE OF INVENTION:  RNA OF MYCOBACTERIA, DERIVED PROBES AND PRIMERS, REAGENT
;    TITLE OF INVENTION:  AND DETECTION METHOD
;    NUMBER OF SEQUENCES:  178
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  OLIFF & BERRIDGE
;      STREET:  700 South Washington Street, Suite 300
;      CITY:  Alexandria
;      STATE:  Virginia
;      COUNTRY:  USA
;      ZIP:  22314
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
```

```
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/403,762A
;       FILING DATE:  23-MAR-1995
;       CLASSIFICATION:  536
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Berridge, William P
;       REGISTRATION NUMBER:  30,024
;       REFERENCE/DOCKET NUMBER:  WPB 29658
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  703-836-6400
;       TELEFAX:  703-836-2787
;   INFORMATION FOR SEQ ID NO:  172:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  98 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  rRNA
;     ORIGINAL SOURCE:
;       ORGANISM:  M. SIMIAE
;       STRAIN:  ATCC 25275
;     POSITION IN GENOME:
;       MAP POSITION:  1409..1420, with respect to the numbering of
;       MAP POSITION:  E. coli
US-08-403-762A-172

  Query Match             66.0%;  Score 13.2;  DB 1;  Length 98;
  Best Local Similarity   83.3%;  Pred. No. 4e+02;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 TGCTTCCCACCCTGAATG 18
          |||||| ||||| ||| |
Db     56 TGCTTCGCACCCCGAAGG 39


RESULT 5
US-09-496-632C-8/c
; Sequence 8, Application US/09496632C
; Patent No. 6468789
; GENERAL INFORMATION:
;  APPLICANT: BAYSAL, Bora E.
;  APPLICANT:  FERRELL, Robert E.
;  APPLICANT:  DEVLIN, Bernie J.
;  APPLICANT:  WILLETT-BROZICK, Joan E.
;  TITLE OF INVENTION: OXYGEN SENSING AND HYPOXIC SELECTION FOR TUMORS
;  FILE REFERENCE: 99-484-US
;  CURRENT APPLICATION NUMBER: US/09/496,632C
;  CURRENT FILING DATE:  2000-02-02
;  NUMBER OF SEQ ID NOS: 18
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
;   LENGTH: 19
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-496-632C-8

  Query Match             65.0%;  Score 13;  DB 4;  Length 19;
  Best Local Similarity   100.0%;  Pred. No. 3.9e+02;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      3 CTTCCCACCCTGA 15
          |||||||||||||
Db     13 CTTCCCACCCTGA 1


RESULT 6
US-08-481-003-21
; Sequence 21, Application US/08481003
; Patent No. 5741899
;  GENERAL INFORMATION:
;    APPLICANT:  CAPON, DANIEL J
;    APPLICANT:  TIAN, HUAN
;    APPLICANT:  SMITH, DOUGLAS H
;    APPLICANT:  WINSLOW, GENINE A
;    APPLICANT:  SIEKEVITZ, MIRIAM
;    TITLE OF INVENTION:  CHIMERIC RECEPTORS FOR REGULATING
;    TITLE OF INVENTION:  CELLULAR PROLIFERATION AND EFFECTOR FUNCTION
;    NUMBER OF SEQUENCES:  31
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  CELL GENESYS, INC.
;      STREET:  322 LAKESIDE DRIVE
;      CITY:  FOSTER CITY
;      STATE:  CALIFORNIA
;      COUNTRY:  USA
;      ZIP:  94404
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/481,003
;      FILING DATE:
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/382,846
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  KRUPEN, KAREN I
;      REGISTRATION NUMBER:  34,647
;      REFERENCE/DOCKET NUMBER:  CELL 17
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 358-9600 x131
;      TELEFAX:  (415) 349-7392
;  INFORMATION FOR SEQ ID NO:  21:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  49 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-481-003-21

  Query Match             65.0%;  Score 13;  DB 1;  Length 49;
  Best Local Similarity   100.0%;  Pred. No. 4.5e+02;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      4 TTCCCACCCTGAA 16
          |||||||||||||
Db     33 TTCCCACCCTGAA 45


RESULT 7
US-08-485-598-21
; Sequence 21, Application US/08485598
; Patent No. 6077947
;  GENERAL INFORMATION:
;    APPLICANT:  CAPON, DANIEL J
;    APPLICANT:  TIAN, HUAN
;    APPLICANT:  SMITH, DOUGLAS H
;    APPLICANT:  WINSLOW, GENINE A
;    APPLICANT:  SIEKEVITZ, MIRIAM
;    TITLE OF INVENTION:  CHIMERIC RECEPTORS FOR REGULATING
;    TITLE OF INVENTION:  CELLULAR PROLIFERATION AND EFFECTOR FUNCTION
;    NUMBER OF SEQUENCES:  31
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  CELL GENESYS, INC.
;      STREET:  322 LAKESIDE DRIVE
;      CITY:  FOSTER CITY
;      STATE:  CALIFORNIA
;      COUNTRY:  USA
;      ZIP:  94404
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
```

```
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/485,598
;      FILING DATE:
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/382,846
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  KRUPEN, KAREN I
;      REGISTRATION NUMBER:  34,647
;      REFERENCE/DOCKET NUMBER:  CELL 17
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 358-9600 x131
;      TELEFAX:  (415) 349-7392
;  INFORMATION FOR SEQ ID NO:  21:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  49 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-485-598-21

  Query Match             65.0%;  Score 13;  DB 3;  Length 49;
  Best Local Similarity   100.0%;  Pred. No. 4.5e+02;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       4 TTCCCACCCTGAA 16
           |||||||||||||
Db      33 TTCCCACCCTGAA 45


RESULT 8
US-07-808-455A-4
; Sequence 4, Application US/07808455A
; Patent No. 5405745
;  GENERAL INFORMATION:
;    APPLICANT:  Gorman, Jessica A.
;    APPLICANT:  Bingham, Catherine
;    TITLE OF INVENTION:  METHOD FOR DETECTING CANDIDA ALBICANS
;    NUMBER OF SEQUENCES:  6
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Burton Rodney
;      STREET:  P.O. Box 4000
;      CITY:  Princeton
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  08543-4000
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/808,455A
;      FILING DATE:  19911217
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Bogden, James M.
;      REGISTRATION NUMBER:  32,962
;      REFERENCE/DOCKET NUMBER:  GP65
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (609) 921-4163
;      TELEFAX:  (609) 921-4526
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  21 base pairs
;      TYPE:  NUCLEIC ACID
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-07-808-455A-4

  Query Match             64.0%;  Score 12.8;  DB 1;  Length 21;
  Best Local Similarity   87.5%;  Pred. No. 5e+02;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       2 GCTTCCCACCCTGAAT 17
           |||||||||||| | ||
Db       5 GCTTCCCACCCGGGAT 20


RESULT 9
US-08-403-888A-71/c
; Sequence 71, Application US/08403888A
; Patent No. 5952490
;  GENERAL INFORMATION:
;    APPLICANT:  Hanecak et al.
;    TITLE OF INVENTION:  Oligonucleotides Having A Conserved G4 Core
;    TITLE OF INVENTION:  Sequence
;    NUMBER OF SEQUENCES:  146
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Woodcock Washburn Kurtz Mackiewicz & No. 5952490ris LLP
;      STREET:  One Liberty Place - 46th Floor
;      CITY:  Philadelphia
;      STATE:  PA
;      COUNTRY:  U.S.A.
;      ZIP:  19103
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5 inch disk, 1.44 Mb
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  WordPerfect 6.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/403,888A
;      FILING DATE:  12-JUN-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/954,185
;      FILING DATE:  29-SEP-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Paul K. Legaard
;      REGISTRATION NUMBER:  38,534
;      REFERENCE/DOCKET NUMBER:  ISIS-1229
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  215-568-3100
;      TELEFAX:  215-568-3439
;  INFORMATION FOR SEQ ID NO:  71:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  21
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-403-888A-71

  Query Match             64.0%;  Score 12.8;  DB 2;  Length 21;
  Best Local Similarity   87.5%;  Pred. No. 5e+02;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       4 TTCCCACCCTGAATGA 19
           | |||||||  ||||||
Db      18 TACCCACCCAGAATGA 3


RESULT 10
US-08-430-225A-11
; Sequence 11, Application US/08430225A
; Patent No. 6204000
;  GENERAL INFORMATION:
;    APPLICANT:  Dong, Jin-Tang; Barrett,
;    APPLICANT:  J. Carl; Lamb, Patricia W.; Isaacs, John T.
;    TITLE OF INVENTION:  DIAGNOSTIC METHODS AND
;    TITLE OF INVENTION:  GENE THERAPY USING REAGENTS DERIVED FROM THE
```

```
;    TITLE OF INVENTION:  HUMAN METASTASIS SUPPRESSOR GENE KAI1
;    NUMBER OF SEQUENCES:  20
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORGAN & FINNEGAN, L.L.P.
;      STREET:  345 PARK AVENUE
;      CITY:  NEW YORK
;      STATE:  NEW YORK
;      COUNTRY:  USA
;      ZIP:  10154
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  FLOPPY DISK
;      COMPUTER:  IBM PC COMPATIBLE
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  WORDPERFECT 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/430,225A
;      FILING DATE:  28-APR-1995
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  RICHARD W. BORK
;      REGISTRATION NUMBER:  36,459
;      REFERENCE/DOCKET NUMBER:  2026-4172
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 758-4800
;      TELEFAX:  (212) 751-6849
;      TELEX:  421792
;  INFORMATION FOR SEQ ID NO:  11:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  21 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-430-225A-11

  Query Match             64.0%;  Score 12.8;  DB 4;  Length 21;
  Best Local Similarity   87.5%;  Pred. No. 5e+02;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      4 TTCCCACCCTGAATGA 19
          | |||||||||| |||
Db      6 TGCCCACCCTGACTGA 21


RESULT 11
US-08-430-225A-14/c
; Sequence 14, Application US/08430225A
; Patent No. 6204000
;  GENERAL INFORMATION:
;    APPLICANT:  Dong, Jin-Tang; Barrett,
;    APPLICANT:  J. Carl; Lamb, Patricia W.; Isaacs, John T.
;    TITLE OF INVENTION:  DIAGNOSTIC METHODS AND
;    TITLE OF INVENTION:  GENE THERAPY USING REAGENTS DERIVED FROM THE
;    TITLE OF INVENTION:  HUMAN METASTASIS SUPPRESSOR GENE KAI1
;    NUMBER OF SEQUENCES:  20
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORGAN & FINNEGAN, L.L.P.
;      STREET:  345 PARK AVENUE
;      CITY:  NEW YORK
;      STATE:  NEW YORK
;      COUNTRY:  USA
;      ZIP:  10154
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  FLOPPY DISK
;      COMPUTER:  IBM PC COMPATIBLE
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  WORDPERFECT 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/430,225A
;      FILING DATE:  28-APR-1995
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  RICHARD W. BORK
;      REGISTRATION NUMBER:  36,459
;      REFERENCE/DOCKET NUMBER:  2026-4172
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 758-4800
;      TELEFAX:  (212) 751-6849
;      TELEX:  421792
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  21 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-430-225A-14

  Query Match             64.0%;  Score 12.8;  DB 4;  Length 21;
  Best Local Similarity   87.5%;  Pred. No. 5e+02;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      4 TTCCCACCCTGAATGA 19
          | |||||||||| |||
Db     16 TGCCCACCCTGACTGA 1


RESULT 12
US-08-790-963-43
; Sequence 43, Application US/08790963
; Patent No. 5837464
;  GENERAL INFORMATION:
;    APPLICANT:  Daniel J. Capon
;    APPLICANT:  Christos John Petropoulous
;    TITLE OF INVENTION:  Compositions And Methods For
;    TITLE OF INVENTION:  Determining Anti-viral Drug Susceptibility And
;    TITLE OF INVENTION:  Resistance
;    NUMBER OF SEQUENCES:  105
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Cooper & Dunham LLP
;      STREET:  1185 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  United States
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  Patentin Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/790,963
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  White, John P.
;      REGISTRATION NUMBER:  28,678
;      REFERENCE/DOCKET NUMBER:  50130-B/JPW/AKC
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  212-278-0400
;      TELEFAX:  212-391-0526
;  INFORMATION FOR SEQ ID NO:  43:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  35 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-790-963-43

  Query Match             64.0%;  Score 12.8;  DB 2;  Length 35;
  Best Local Similarity   87.5%;  Pred. No. 5.4e+02;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      4 TTCCCACCCTGAATGA 19
          ||||||| ||| |||
Db     10 TTCCCACCATGATTGA 25
```

```
RESULT 13
US-09-371-774-43
; Sequence 43, Application US/09371774
; Patent No. 6242187
;   GENERAL INFORMATION:
;        APPLICANT: Daniel J. Capon
;                   Christos John Petropoulous
;        TITLE OF INVENTION: Compositions And Methods For
;                            Determining Anti-viral Drug Susceptibility And
;                            Resistance and Anti-viral Drug Screening
;        NUMBER OF SEQUENCES: 105
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Cooper & Dunham LLP
;             STREET: 1185 Avenue of the Americas
;             CITY: New York
;             STATE: New York
;             COUNTRY: United States
;             ZIP: 10036
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: Patentin Release #1.0, Version#1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/371,774
;             FILING DATE: 10-Aug-1999
;             CLASSIFICATION: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: White, John P.
;             REGISTRATION NUMBER: 28,678
;             REFERENCE/DOCKET NUMBER: 50130-F/JPW/CMR
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 212-278-0400
;             TELEFAX: 212-391-0526
;   INFORMATION FOR SEQ ID NO: 43:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 35 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: DNA (genomic)
;        SEQUENCE DESCRIPTION: SEQ ID NO: 43:
US-09-371-774-43

  Query Match             64.0%;  Score 12.8;  DB 4;  Length 35;
  Best Local Similarity   87.5%;  Pred. No. 5.4e+02;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       4 TTCCCACCCTGAATGA 19
           |||||||| ||| |||
Db      10 TTCCCACCATGATTGA 25


RESULT 14
US-08-518-878B-3
; Sequence 3, Application US/08518878B
; Patent No. 5702902
;  GENERAL INFORMATION:
;    APPLICANT:  Tartaglia, Louis A.
;    TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THE
;    TITLE OF INVENTION:  TREATMENT OF BODY WEIGHT DISORDERS, INCLUDING OBESITY
;    NUMBER OF SEQUENCES:  57
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  U.S.A.
;      ZIP:  10036-2711
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/518,878B
;      FILING DATE:  23-AUG-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Coruzzi, Laura A.
;      REGISTRATION NUMBER:  30,742
;      REFERENCE/DOCKET NUMBER:  7853-036
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 790-9090
;      TELEFAX:  (212) 869-9741/8864
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  95 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-518-878B-3

  Query Match             64.0%;  Score 12.8;  DB 1;  Length 95;
  Best Local Similarity   87.5%;  Pred. No. 6.4e+02;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       4 TTCCCACCCTGAATGA 19
           |||||  |||||||||
Db      52 TTCCCTTCCTGAATGA 67


RESULT 15
US-08-294-522B-3
; Sequence 3, Application US/08294522B
; Patent No. 5741666
;  GENERAL INFORMATION:
;    APPLICANT:  Tartaglia, Louis A.
;    TITLE OF INVENTION:  Compositions and Methods for the
;    TITLE OF INVENTION:  Treatment of Body Weight Disorders, Including Obesity
;    NUMBER OF SEQUENCES:  48
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Pennie & Edmonds
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      ZIP:  10036-2711
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/294,522B
;      FILING DATE:  23-AUG-1994
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Coruzzi, Laura A.
;      REGISTRATION NUMBER:  30,742
;      REFERENCE/DOCKET NUMBER:  7853-015
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 790-9090
;      TELEFAX:  (212) 869-8864/9741
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  95 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  both
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
US-08-294-522B-3
```

Query Match 64.0%; Score 12.8; DB 1; Length 95;
Best Local Similarity 87.5%; Pred. No. 6.4e+02;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy       4 TTCCCACCCTGAATGA 19
           |||||  |||||||||
Db      52 TTCCCTTCCTGAATGA 67
```

Search completed: November 23, 2002, 06:36:16
Job time : 22.55 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 05:54:41 ; Search time 17.25 Seconds
(without alignments)
439.108 Million cell updates/sec

Title: US-09-296-264-14
Perfect score: 20
Sequence: 1 tgcttcccaccctgaatgat 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 335578 seqs, 189365133 residues

Total number of hits satisfying chosen parameters: 195614

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	13.2	66.0	30	10	US-09-911-777-10	Sequence 10, Appl
	2	13.2	66.0	99	10	US-09-864-761-28802	Sequence 28802, A
	3	12.8	64.0	21	10	US-09-795-380-11	Sequence 11, Appl
c	4	12.8	64.0	21	10	US-09-795-380-14	Sequence 14, Appl
c	5	12.8	64.0	91	10	US-09-864-761-19364	Sequence 19364, A
	6	12.6	63.0	32	9	US-09-779-050A-50	Sequence 50, Appl
	7	12.6	63.0	34	10	US-09-725-285-5	Sequence 5, Appli
	8	12.6	63.0	34	10	US-09-779-879A-5	Sequence 5, Appli
	9	12.6	63.0	34	10	US-09-779-880A-5	Sequence 5, Appli
	10	12.6	63.0	34	10	US-09-195-662A-5	Sequence 5, Appli
	11	12.6	63.0	34	10	US-09-339-912A-5	Sequence 5, Appli
	12	12.6	63.0	34	10	US-09-502-783A-5	Sequence 5, Appli
	13	12.6	63.0	35	10	US-09-867-569-5	Sequence 5, Appli
	14	12.6	63.0	66	10	US-09-983-965-2148	Sequence 2148, Ap
c	15	12.6	63.0	78	10	US-09-972-809-6	Sequence 6, Appli
	16	12.6	63.0	89	10	US-09-764-869-2114	Sequence 2114, Ap
	17	12.6	63.0	89	10	US-09-764-877-3966	Sequence 3966, Ap
	18	12.6	63.0	94	10	US-09-895-828-204	Sequence 204, App
	19	12.4	62.0	28	8	US-08-591-486B-116	Sequence 116, App
	20	12.4	62.0	61	10	US-09-973-451-34	Sequence 34, Appl
	21	12.2	61.0	31	10	US-09-801-274-622	Sequence 622, App
	22	12.2	61.0	42	10	US-09-971-798-24	Sequence 24, Appl
	23	12.2	61.0	45	10	US-09-987-485-4	Sequence 4, Appli
c	24	12.2	61.0	48	10	US-09-971-798-25	Sequence 25, Appl
	25	12.2	61.0	51	10	US-09-904-599A-5	Sequence 5, Appli
	26	12.2	61.0	63	10	US-09-971-798-12	Sequence 12, Appl
c	27	12.2	61.0	63	10	US-09-971-798-13	Sequence 13, Appl
	28	12.2	61.0	90	10	US-09-917-330-2	Sequence 2, Appli
	29	12.2	61.0	90	10	US-09-917-330-3	Sequence 3, Appli
	30	12.2	61.0	98	10	US-09-294-093B-4749	Sequence 4749, Ap
c	31	12	60.0	28	10	US-09-949-145-57	Sequence 57, Appl
	32	12	60.0	34	10	US-09-810-796-16	Sequence 16, Appl
	33	12	60.0	57	12	US-10-124-796-1	Sequence 1, Appli
c	34	12	60.0	85	10	US-09-864-761-29376	Sequence 29376, A
c	35	12	60.0	91	10	US-09-864-761-18106	Sequence 18106, A
	36	12	60.0	93	10	US-09-294-093B-913	Sequence 913, App
c	37	11.8	59.0	31	10	US-09-801-274-1174	Sequence 1174, Ap
	38	11.8	59.0	34	12	US-10-084-206-6	Sequence 6, Appli
c	39	11.8	59.0	66	10	US-09-783-590-9266	Sequence 9266, Ap
	40	11.6	58.0	22	12	US-10-139-262-21	Sequence 21, Appl
	41	11.6	58.0	24	10	US-09-846-808-40	Sequence 40, Appl
	42	11.6	58.0	27	10	US-09-027-287-8	Sequence 8, Appli
	43	11.6	58.0	27	10	US-09-027-287-10	Sequence 10, Appl
	44	11.6	58.0	27	10	US-09-027-287-19	Sequence 19, Appl
	45	11.6	58.0	27	10	US-09-252-656B-8	Sequence 8, Appli

ALIGNMENTS

RESULT 1
US-09-911-777-10
; Sequence 10, Application US/09911777
; Patent No. US20020037852A1
; GENERAL INFORMATION:
; APPLICANT: BIOGEN, INC.
; APPLICANT: APOTECH S.A.
; APPLICANT: BROWNING, Jeffrey
; APPLICANT: AMBROSE, Christine
; APPLICANT: MACKAY, Fabienne
; APPLICANT: TSCHOPP, Jurg
; APPLICANT: SCHNEIDER, Pascal
; TITLE OF INVENTION: BAFF, Inhibitors Thereof and Their Use
; TITLE OF INVENTION: in the Modulation of B-Cell Response
; FILE REFERENCE: A070 US
; CURRENT APPLICATION NUMBER: US/09/911,777
; CURRENT FILING DATE: 2001-07-24
; PRIOR APPLICATION NUMBER: 60/117,169
; PRIOR FILING DATE: 1999-01-25
; PRIOR APPLICATION NUMBER: 60/143,228
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: PCT/US00/01788
; PRIOR FILING DATE: 2000-01-25
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 30
; TYPE: DNA
; ORGANISM: Homo Sapien
US-09-911-777-10

Query Match 66.0%; Score 13.2; DB 10; Length 30;
Best Local Similarity 83.3%; Pred. No. 7.6e+02;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy       2 GCTTCCCACCCTGAATGA 19
           |||| ||||| || ||||
Db       6 GCTTGCCACCATGGATGA 23
```

RESULT 2
US-09-864-761-28802

```
; Sequence 28802, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 28802
;   LENGTH: 99
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AC007492.2
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.95
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.87
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.2
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.75
;   OTHER INFORMATION: EST_HUMAN HIT: BE175575.1, EVALUE 2.00e-08
;   OTHER INFORMATION: NT HIT: AB042815.1, EVALUE 4.00e-12
US-09-864-761-28802

  Query Match             66.0%;  Score 13.2;  DB 10;  Length 99;
  Best Local Similarity   83.3%;  Pred. No. 8.7e+02;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 GCTTCCCACCCTGAATGA 19
           | || ||||| |||||||
Db      30 GGTTTCCACCATGAATGA 47
```

```
RESULT 3
US-09-795-380-11
; Sequence 11, Application US/09795380
; Patent No. US20020058257A1
;   GENERAL INFORMATION:
;        APPLICANT: Dong, Jin-Tang; Barrett,
;                   J. Carl; Lamb, Patricia W.; Isaacs, John T.
;        TITLE OF INVENTION: DIAGNOSTIC METHODS AND
;                            GENE THERAPY USING REAGENTS DERIVED FROM THE
;                            HUMAN METASTASIS SUPPRESSOR GENE KAI1
;        NUMBER OF SEQUENCES: 20
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
;             STREET: 345 PARK AVENUE
;             CITY: NEW YORK
;             STATE: NEW YORK
;             COUNTRY: USA
;             ZIP: 10154
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: FLOPPY DISK
;             COMPUTER: IBM PC COMPATIBLE
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: MICROSOFT WORD 97
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/795,380
;             FILING DATE: 27-Feb-2001
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 09/232,507
;             FILING DATE: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: RICHARD W. BORK
;             REGISTRATION NUMBER: 36,459
;             REFERENCE/DOCKET NUMBER: 2026-4172US1
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (212) 758-4800
;             TELEFAX: (212) 751-6849
;             TELEX: 421792
;   INFORMATION FOR SEQ ID NO: 11:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 21 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        SEQUENCE DESCRIPTION: SEQ ID NO: 11:
US-09-795-380-11

  Query Match             64.0%;  Score 12.8;  DB 10;  Length 21;
  Best Local Similarity   87.5%;  Pred. No. 1.2e+03;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       4 TTCCCACCCTGAATGA 19
           | |||||||||| |||
Db       6 TGCCCACCCTGACTGA 21


RESULT 4
US-09-795-380-14/c
; Sequence 14, Application US/09795380
; Patent No. US20020058257A1
;   GENERAL INFORMATION:
;        APPLICANT: Dong, Jin-Tang; Barrett,
;                   J. Carl; Lamb, Patricia W.; Isaacs, John T.
;        TITLE OF INVENTION: DIAGNOSTIC METHODS AND
;                            GENE THERAPY USING REAGENTS DERIVED FROM THE
;                            HUMAN METASTASIS SUPPRESSOR GENE KAI1
;        NUMBER OF SEQUENCES: 20
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
;             STREET: 345 PARK AVENUE
;             CITY: NEW YORK
;             STATE: NEW YORK
;             COUNTRY: USA
;             ZIP: 10154
```

```
;         COMPUTER READABLE FORM:
;              MEDIUM TYPE: FLOPPY DISK
;              COMPUTER: IBM PC COMPATIBLE
;              OPERATING SYSTEM: PC-DOS/MS-DOS
;              SOFTWARE: MICROSOFT WORD 97
;         CURRENT APPLICATION DATA:
;              APPLICATION NUMBER: US/09/795,380
;              FILING DATE: 27-Feb-2001
;         PRIOR APPLICATION DATA:
;              APPLICATION NUMBER: 09/232,507
;              FILING DATE: <Unknown>
;         ATTORNEY/AGENT INFORMATION:
;              NAME: RICHARD W. BORK
;              REGISTRATION NUMBER: 36,459
;              REFERENCE/DOCKET NUMBER: 2026-4172US1
;         TELECOMMUNICATION INFORMATION:
;              TELEPHONE: (212) 758-4800
;              TELEFAX: (212) 751-6849
;              TELEX: 421792
;    INFORMATION FOR SEQ ID NO: 14:
;         SEQUENCE CHARACTERISTICS:
;              LENGTH: 21 base pairs
;              TYPE: nucleic acid
;              STRANDEDNESS: single
;              TOPOLOGY: linear
;         SEQUENCE DESCRIPTION: SEQ ID NO: 14:
US-09-795-380-14

  Query Match             64.0%;  Score 12.8;  DB 10;  Length 21;
  Best Local Similarity   87.5%;  Pred. No. 1.2e+03;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       4 TTCCCACCCTGAATGA 19
           | |||||||||| |||
Db      16 TGCCCACCCTGACTGA 1


RESULT 5
US-09-864-761-19364/c
; Sequence 19364, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 19364
;   LENGTH: 91
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AC009411.1
;   OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.1
;   OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 0.97
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.1
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL =   1
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.2
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.9
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.1
;   OTHER INFORMATION: NT HIT: K01871.1, EVALUE 3.00e+00
;   OTHER INFORMATION: EST_HUMAN HIT: BF686309.1, EVALUE 2.80e-01
US-09-864-761-19364

  Query Match             64.0%;  Score 12.8;  DB 10;  Length 91;
  Best Local Similarity   87.5%;  Pred. No. 1.4e+03;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 TGCTTCCCACCCTGAA 16
           || || ||||||||||
Db      84 TGATTTCCACCCTGAA 69


RESULT 6
US-09-779-050A-50
; Sequence 50, Application US/09779050A
; Patent No. US20020160416A1
; GENERAL INFORMATION:
;  APPLICANT: BOYLE, WILLIAM
;  APPLICANT:  HSU, HAILING
;  TITLE OF INVENTION: RECEPTOR FROM TNF FAMILY
;  FILE REFERENCE: A-570B
;  CURRENT APPLICATION NUMBER: US/09/779,050A
;  CURRENT FILING DATE:  2001-02-12
;  PRIOR APPLICATION NUMBER: 60/181,800
;  PRIOR FILING DATE: 2000-02-11
;  NUMBER OF SEQ ID NOS: 52
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 50
;   LENGTH: 32
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-779-050A-50

  Query Match             63.0%;  Score 12.6;  DB 9;  Length 32;
  Best Local Similarity   78.9%;  Pred. No. 1.5e+03;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 GCTTCCCACCCTGAATGAT 20
           |||||| | ||||| | ||
Db       8 GCTTCCGATCCTGAGTAAT 26


RESULT 7
```

```
US-09-725-285-5
; Sequence 5, Application US/09725285
; Patent No. US20010000241A1
; GENERAL INFORMATION:
;  APPLICANT:   Li, Yi
;  APPLICANT:  Ruben, Steven, M.
;  TITLE OF INVENTION:   Antibodies to Human G-Protein Chemokine Receptor HDGNR10
;  TITLE OF INVENTION:  (CCR5 Receptor)
;  FILE REFERENCE:   1488.1150003
;  CURRENT APPLICATION NUMBER: US/09/725,285
;  CURRENT FILING DATE:  2000-11-29
;  PRIOR APPLICATION NUMBER:   09/339,912
;  PRIOR FILING DATE:   1999-06-25
;  PRIOR APPLICATION NUMBER:   09/195,662
;  PRIOR FILING DATE:   1998-11-18
;  PRIOR APPLICATION NUMBER:   08/466,343
;  PRIOR FILING DATE:   1995-06-06
;  NUMBER OF SEQ ID NOS:   9
;  SOFTWARE:   PatentIn version 3.0
; SEQ ID NO 5
;   LENGTH: 34
;   TYPE: DNA
;   ORGANISM: Oligonucleotide
US-09-725-285-5

  Query Match             63.0%;  Score 12.6;  DB 10;  Length 34;
  Best Local Similarity   78.9%;  Pred. No. 1.5e+03;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      2 GCTTCCCACCCTGAATGAT 20
          |||| ||||| || || ||
Db      7 GCTTGCCACCATGGATTAT 25


RESULT 8
US-09-779-879A-5
; Sequence 5, Application US/09779879A
; Patent No. US20020048786A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen, Craig A.
;  APPLICANT:  Roschke, Viktor
;  APPLICANT:  Li, Yi
;  APPLICANT:  Ruben, Steven, M.
;  TITLE OF INVENTION: Human G-protein Chemokine Receptor (CCR5) HDGNR10
;  FILE REFERENCE: 1488.115000A
;  CURRENT APPLICATION NUMBER: US/09/779,879A
;  CURRENT FILING DATE:  2001-02-09
;  PRIOR APPLICATION NUMBER: US 60/181,258
;  PRIOR FILING DATE: 2000-02-09
;  PRIOR APPLICATION NUMBER: US 60/187,999
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: US 60/234,336
;  PRIOR FILING DATE: 2000-09-22
;  NUMBER OF SEQ ID NOS: 58
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
;   LENGTH: 34
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: 5' Oligonucleotide primer for HDGNR10
US-09-779-879A-5

  Query Match             63.0%;  Score 12.6;  DB 10;  Length 34;
  Best Local Similarity   78.9%;  Pred. No. 1.5e+03;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      2 GCTTCCCACCCTGAATGAT 20
          |||| ||||| || || ||
Db      7 GCTTGCCACCATGGATTAT 25


RESULT 9
US-09-779-880A-5
; Sequence 5, Application US/09779880A
; Patent No. US20020061834A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen, Craig A.
;  APPLICANT:  Roschke, Viktor
;  APPLICANT:  Li, Yi
;  APPLICANT:  Ruben, Steven, M.
;  TITLE OF INVENTION: Human G-protein Chemokine Receptor (CCR5) HDGNR10
;  FILE REFERENCE: 1488.115000C
;  CURRENT APPLICATION NUMBER: US/09/779,880A
;  CURRENT FILING DATE:  2001-02-09
;  PRIOR APPLICATION NUMBER: US 60/181,258
;  PRIOR FILING DATE: 2000-02-09
;  PRIOR APPLICATION NUMBER: US 60/187,999
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: US 60/234,336
;  PRIOR FILING DATE: 2000-09-22
;  NUMBER OF SEQ ID NOS: 58
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
;   LENGTH: 34
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: 5' Oligonucleotide primer for HDGNR10
US-09-779-880A-5

  Query Match             63.0%;  Score 12.6;  DB 10;  Length 34;
  Best Local Similarity   78.9%;  Pred. No. 1.5e+03;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      2 GCTTCCCACCCTGAATGAT 20
          |||| ||||| || || ||
Db      7 GCTTGCCACCATGGATTAT 25


RESULT 10
US-09-195-662A-5
; Sequence 5, Application US/09195662A
; Patent No. US20020076745A1
; GENERAL INFORMATION:
;  APPLICANT:   Li, Yi
;  APPLICANT:  Ruben, Steven, M.
;  TITLE OF INVENTION:   Human G-Protein Chemokine Receptor HDGNR10 (CCR5 Receptor)
;  FILE REFERENCE:   1488.1150002
;  CURRENT APPLICATION NUMBER: US/09/195,662A
;  CURRENT FILING DATE:  1998-11-18
;  PRIOR APPLICATION NUMBER:   08/466,343
;  PRIOR FILING DATE:   1995-06-06
;  NUMBER OF SEQ ID NOS:   9
;  SOFTWARE:   PatentIn version 3.0
; SEQ ID NO 5
;   LENGTH: 34
;   TYPE: DNA
;   ORGANISM: Oligonucleotide
US-09-195-662A-5

  Query Match             63.0%;  Score 12.6;  DB 10;  Length 34;
  Best Local Similarity   78.9%;  Pred. No. 1.5e+03;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      2 GCTTCCCACCCTGAATGAT 20
          |||| ||||| || || ||
Db      7 GCTTGCCACCATGGATTAT 25


RESULT 11
US-09-339-912A-5
; Sequence 5, Application US/09339912A
; Patent No. US20020099176A1
; GENERAL INFORMATION:
;  APPLICANT:   Li, Yi
```

```
;  APPLICANT:  Ruben, Steven, M.
;  TITLE OF INVENTION:   Antibodies to Human G-Protein Chemokine Receptor HDGNR10
;  TITLE OF INVENTION:  (CCR5 Receptor)
;  FILE REFERENCE:   1488.1150003
;  CURRENT APPLICATION NUMBER: US/09/339,912A
;  CURRENT FILING DATE:  1999-06-25
;  PRIOR APPLICATION NUMBER:   09/195,662
;  PRIOR FILING DATE:   1998-11-18
;  PRIOR APPLICATION NUMBER:   08/466,343
;  PRIOR FILING DATE:   1995-06-06
;  NUMBER OF SEQ ID NOS:   9
;  SOFTWARE:   PatentIn version 3.0
; SEQ ID NO 5
;   LENGTH: 34
;   TYPE: DNA
;   ORGANISM: Oligonucleotide
US-09-339-912A-5

  Query Match            63.0%;  Score 12.6;  DB 10;  Length 34;
  Best Local Similarity  78.9%;  Pred. No. 1.5e+03;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      2 GCTTCCCACCCTGAATGAT 20
          |||| ||||| || || ||
Db      7 GCTTGCCACCATGGATTAT 25


RESULT 12
US-09-502-783A-5
; Sequence 5, Application US/09502783A
; Patent No. US20020132269A1
; GENERAL INFORMATION:
;  APPLICANT: Li, Yi
;  APPLICANT:  Ruben, Steven M.
;  TITLE OF INVENTION: Polynucleotides Encoding Human G-Protein Chemokine Receptor (CCR5
;  TITLE OF INVENTION:  HDGNR10
;  FILE REFERENCE: 1488.1150006
;  CURRENT APPLICATION NUMBER: US/09/502,783A
;  CURRENT FILING DATE:  2001-08-23
;  PRIOR APPLICATION NUMBER: 08/466,343
;  PRIOR FILING DATE: 1995-06-06
;  NUMBER OF SEQ ID NOS: 9
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
;   LENGTH: 34
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Oligonucleotide
US-09-502-783A-5

  Query Match            63.0%;  Score 12.6;  DB 10;  Length 34;
  Best Local Similarity  78.9%;  Pred. No. 1.5e+03;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      2 GCTTCCCACCCTGAATGAT 20
          |||| ||||| || || ||
Db      7 GCTTGCCACCATGGATTAT 25


RESULT 13
US-09-867-569-5
; Sequence 5, Application US/09867569
; Patent No. US20010036650A1
; GENERAL INFORMATION:
;  APPLICANT: Li et al.
;  TITLE OF INVENTION: C5a Receptor
;  FILE REFERENCE: PF130D1C1
;  CURRENT APPLICATION NUMBER: US/09/867,569
;  CURRENT FILING DATE:  2001-05-31
;  PRIOR APPLICATION NUMBER: 09/082,529
;  PRIOR FILING DATE: 1998-05-21
;  PRIOR APPLICATION NUMBER: 08/458,970
;  PRIOR FILING DATE: 1995-06-02
;  PRIOR APPLICATION NUMBER: PCT/US94/09234
;  PRIOR FILING DATE: 1994-08-16
;  NUMBER OF SEQ ID NOS: 11
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
;   LENGTH: 35
;   TYPE: DNA
;   ORGANISM: artificial sequence
;   FEATURE:
;   NAME/KEY: primer_bind
;   LOCATION: (1)..(35)
;   OTHER INFORMATION: primer containing a HindIII site followed by 18 nucleotides
;   OTHER INFORMATION: of C5a receptor coding sequence starting from the initiation c
US-09-867-569-5

  Query Match            63.0%;  Score 12.6;  DB 10;  Length 35;
  Best Local Similarity  78.9%;  Pred. No. 1.5e+03;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      2 GCTTCCCACCCTGAATGAT 20
          |||| ||||| || | |||
Db      8 GCTTGCCACCATGGAAGAT 26


RESULT 14
US-09-983-965-2148
; Sequence 2148, Application US/09983965
; Patent No. US20020137160A1
; GENERAL INFORMATION:
;  APPLICANT: Warren, Wesley C.
;  APPLICANT:  Tao, Nengbing
;  APPLICANT:  Byatt, John C.
;  APPLICANT:  Mathialagan, Nagappan
;  TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
;  TITLE OF INVENTION:  MUSCLE AND FAT DEPOSITION
;  FILE REFERENCE: 37-21(10297)C
;  CURRENT APPLICATION NUMBER: US/09/983,965
;  CURRENT FILING DATE:  2001-10-26
;  PRIOR APPLICATION NUMBER: US 09/465,231
;  PRIOR FILING DATE: 1999-12-15
;  PRIOR APPLICATION NUMBER: US 60/113,678
;  PRIOR FILING DATE: 1998-12-17
;  NUMBER OF SEQ ID NOS: 5912
; SEQ ID NO 2148
;   LENGTH: 66
;   TYPE: DNA
;   ORGANISM: Bos taurus
;   FEATURE:
;   OTHER INFORMATION: Clone ID: 38-LIB3057-015-Q1-K1-B6
US-09-983-965-2148

  Query Match            63.0%;  Score 12.6;  DB 10;  Length 66;
  Best Local Similarity  78.9%;  Pred. No. 1.6e+03;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 TGCTTCCCACCCTGAATGA 19
          | ||| |||||||||| | |
Db     47 TCCTTTCCACCCTGACTCA 65


RESULT 15
US-09-972-809-6/c
; Sequence 6, Application US/09972809
; Patent No. US20020151490A1
; GENERAL INFORMATION:
;  APPLICANT: Sundeep, Khosla
;  APPLICANT:  Conover, Cheryl A.
;  TITLE OF INVENTION: TREATMENT OF OSTEOPOROSIS
;  FILE REFERENCE: 07039/183001
;  CURRENT APPLICATION NUMBER: US/09/972,809
;  CURRENT FILING DATE:  2001-10-05
;  PRIOR APPLICATION NUMBER: 09/428,226
```

```
;  PRIOR FILING DATE: 1999-10-27
;  PRIOR APPLICATION NUMBER: 60/045,607
;  PRIOR FILING DATE: 1997-05-05
;  NUMBER OF SEQ ID NOS: 7
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6
;   LENGTH: 78
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Synthetically derived primer
US-09-972-809-6

  Query Match             63.0%;  Score 12.6;  DB 10;  Length 78;
  Best Local Similarity   78.9%;  Pred. No. 1.7e+03;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 GCTTCCCACCCTGAATGAT 20
           |||||  ||| | ||||||
Db      33 GCTTCTAACCGTAAATGAT 15


Search completed: November 23, 2002, 06:42:11
Job time : 18.25 secs
```

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 25, 2002, 09:10:06 ; Search time 755.55 Seconds
(without alignments)
428.707 Million cell updates/sec

Title: US-09-296-264-14
Perfect score: 20
Sequence: 1 tgcttcccaccctgaatgat 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 357874

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	15.2	76.0	94	9	AI766407	AI766407 wh61f01.x
c	2	14.4	72.0	50	9	AU105169	AU105169 AU105169
c	3	14.4	72.0	67	9	AA722633	AA722633 zg86g09.s
	4	14.2	71.0	70	9	AI862114	AI862114 tw37a03.x
c	5	14.2	71.0	96	17	BH847460	BH847460 SALK_0544
	6	13.6	68.0	41	17	AZ595851	AZ595851 1M0408B14
c	7	13.6	68.0	62	17	AZ799423	AZ799423 2M0056G22
	8	13.6	68.0	68	13	BM565783	BM565783 rt02d09.y
	9	13.6	68.0	69	14	T65253	T65253 yc79b02.r1
	10	13.6	68.0	92	12	BG508809	BG508809 sac90f07.
c	11	13.6	68.0	97	9	AA885649	AA885649 oj32c12.s
c	12	13.4	67.0	63	17	AZ478351	AZ478351 1M0298B23
	13	13.4	67.0	70	10	AV970203	AV970203 AV970203
c	14	13.4	67.0	86	17	AZ665595	AZ665595 1M0547E07
c	15	13.4	67.0	87	9	AU076672	AU076672 AU076672
c	16	13.4	67.0	91	9	AA073983	AA073983 mm81b08.r
c	17	13.2	66.0	52	9	AA285400	AA285400 va67h12.r
c	18	13.2	66.0	79	9	AU009126	AU009126 AU009126
c	19	13.2	66.0	81	9	AU009127	AU009127 AU009127
c	20	13.2	66.0	97	10	AW190337	AW190337 xl61b07.x
	21	12.8	64.0	40	14	D45800	D45800 HUMGS03017
c	22	12.8	64.0	50	9	AU105174	AU105174 AU105174
	23	12.8	64.0	52	12	BF643317	BF643317 NF006H07E
c	24	12.8	64.0	61	9	AA736682	AA736682 oa22f05.s
	25	12.8	64.0	66	17	TA146H05Q	AL472759 T. brucei
c	26	12.8	64.0	68	17	AZ771388	AZ771388 1M0573F02
	27	12.8	64.0	70	9	AA156701	AA156701 zl08f05.r
c	28	12.8	64.0	73	12	BG485887	BG485887 SWOv3MCAM
	29	12.8	64.0	79	9	AU076877	AU076877 AU076877
	30	12.8	64.0	87	9	AA270233	AA270233 va65h04.r
	31	12.8	64.0	90	17	AZ567355	AZ567355 233PvG01
	32	12.8	64.0	97	17	AZ349648	AZ349648 1M0086B16
	33	12.8	64.0	100	9	AA937861	AA937861 nw90b10.s
	34	12.8	64.0	100	17	AZ775520	AZ775520 2M0008P08
c	35	12.6	63.0	43	9	AI097835	AI097835 ue40e02.y
	36	12.6	63.0	55	10	AV619150	AV619150 AV619150
c	37	12.6	63.0	61	14	T72346	T72346 yc71h11.s1
	38	12.6	63.0	61	17	AL761648	AL761648 Arabidops
c	39	12.6	63.0	67	9	AA429848	AA429848 zw59c10.r
c	40	12.6	63.0	78	9	AA777324	AA777324 zi92g05.s
c	41	12.6	63.0	84	13	BI312886	BI312886 dai28d07.
c	42	12.6	63.0	88	9	AI759160	AI759160 No.15 pol
c	43	12.6	63.0	89	9	AA120466	AA120466 mp83d09.r
	44	12.6	63.0	90	9	AI508511	AI508511 vb35d02.y
	45	12.6	63.0	90	17	AZ767516	AZ767516 1M0566C17

ALIGNMENTS

RESULT 1
AI766407
LOCUS AI766407 94 bp mRNA linear EST 20-DEC-1999
DEFINITION wh61f01.x1 NCI_CGAP_Kid11 Homo sapiens cDNA clone IMAGE:2385241 3' similar to SW:NU4M_PANTR P03906 NADH-UBIQUINONE OXIDOREDUCTASE CHAIN 4 ;contains MSR1.t1 MSR1 repetitive element ;, mRNA sequence.
ACCESSION AI766407
VERSION AI766407.1 GI:5232916
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 94)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html

```
           Trace considered overall poor quality
           Insert Length: 892   Std Error: 0.00
           Seq primer: -40UP from Gibco
           High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .94
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2385241"
                     /clone_lib="NCI_CGAP_Kid11"
                     /lab_host="DH10B"
                     /note="Organ: kidney; Vector: pT7T3D-Pac (Pharmacia) with
                     a modified polylinker; Site_1: Not I; Site_2: Eco RI;
                     Plasmid DNA from the normalized library NCI_CGAP_Kid3 was
                     prepared, and ss circles were made in vitro. Following HAP
                     purification, this DNA was used as tracer in a subtractive
                     hybridization reaction. The driver was PCR-amplified cDNAs
                     from a pool of 5,000 clones made from the same library
                     (cloneIDs 1322376-1323911, 1456007-1456775, and
                     1500552-1502855). Subtraction by Bento Soares and M.
                     Fatima Bonaldo.  "
BASE COUNT       26 a     15 c     33 g     20 t
ORIGIN

  Query Match             76.0%;  Score 15.2;  DB 9;  Length 94;
  Best Local Similarity   85.0%;  Pred. No. 4.2e+03;
  Matches   17;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 TGCTTCCCACCCTGAATGAT 20
           |||||||   ||||||||||
Db      73 TGCTTCCTCGCCTGAATGAT 92


RESULT 2
AU105169/c
LOCUS       AU105169                 50 bp    mRNA    linear   EST 30-AUG-2001
DEFINITION  AU105169 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
            CAS00536, mRNA sequence.
ACCESSION   AU105169
VERSION     AU105169.1  GI:13554690
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Suzuki,Y., Taira,H., Tsunoda,T., Mizushima-Sugano,J., Sese,J., Hata
            ,H., Ota,T., Isogai,T., Tanaka,T., Morishita,S., Okubo,K., Sakaki
            ,Y., Nakamura,Y., Suyama,A. and Sugano,S.
  TITLE     Diverse transcriptional initiation revealed by fine, large-scale
            mapping of mRNA start sites
  JOURNAL   EMBO Rep. 2 (5), 388-393 (2001)
  MEDLINE   21270072
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp
            Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano
            ,S. Construction and characterization of a full length-enriched and
            a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="CAS00536"
                     /clone_lib="Sugano Homo sapiens cDNA library"
                     /note="Differential display comparison of untreated and
                     dimethylfumarate treated U937 cells"
BASE COUNT       11 a     14 c     19 g      6 t
ORIGIN

  Query Match             72.0%;  Score 14.4;  DB 9;  Length 50;
  Best Local Similarity   93.8%;  Pred. No. 8.4e+03;
  Matches   15;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       4 TTCCCACCCTGAATGA 19
           ||||||| ||||||||
Db      18 TTCCCACTCTGAATGA 3


RESULT 3
AA722633/c
LOCUS       AA722633                 67 bp    mRNA    linear   EST 02-JAN-1998
DEFINITION  zg86g09.s1 Soares_fetal_heart_NbHH19W Homo sapiens cDNA clone
            IMAGE:409504 3' similar to TR:O00244 O00244 COPPER TRANSPORT
            PROTEIN HAH1. ;, mRNA sequence.
ACCESSION   AA722633
VERSION     AA722633.1  GI:2740340
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 67)
  AUTHORS   Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
            Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin
            ,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B.,
            White,Y., Wylie,T., Waterston,R. and Wilson,R.
  TITLE     WashU-NCI human EST Project
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Trace considered overall poor quality
            Possible reversed clone: similarity on wrong strand
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .67
                     /organism="Homo sapiens"
                     /db_xref="GDB:1316289"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:409504"
                     /clone_lib="Soares_fetal_heart_NbHH19W"
                     /sex="unknown"
                     /dev_stage="19 weeks"
                     /lab_host="DH10B (ampicillin resistant)"
                     /note="Organ: heart; Vector: pT7T3D (Pharmacia) with a
                     modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCATCTTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was size selected, ligated to Eco RI
                     adapters (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of a modified pT7T3 vector
                     (Pharmacia). Library went through one round of
                     normalization to a Cot = 5. Library constructed by
                     M.Fatima Bonaldo. This library was constructed from the
                     same fetus as the fetal lung library, Soares fetal lung
                     NbHL19W."
BASE COUNT       15 a     14 c     20 g     18 t
ORIGIN

  Query Match             72.0%;  Score 14.4;  DB 9;  Length 67;
  Best Local Similarity   93.8%;  Pred. No. 9.2e+03;
  Matches   15;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 TGCTTCCCACCCTGAA 16
           ||||||| ||||||||
Db      51 TGCTTCCAACCCTGAA 36
```

```
RESULT 4
AI862114
LOCUS       AI862114                 70 bp    mRNA    linear   EST 15-DEC-1999
DEFINITION  tw37a03.x1 NCI_CGAP_Ut1 Homo sapiens cDNA clone IMAGE:2261836 3'
            similar to TR:O61603 O61603 EYELID. ;contains element MSR1
            repetitive element ;, mRNA sequence.
ACCESSION   AI862114
VERSION     AI862114.1  GI:5526221
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 70)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html

            Trace considered overall poor quality
            Insert Length: 927   Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .70
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2261836"
                     /clone_lib="NCI_CGAP_Ut1"
                     /tissue_type="well-differentiated endometrial
                     adenocarcinoma, 7 pooled tumors"
                     /lab_host="DH10B"
                     /note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.75 kb. Life Technologies catalog #:
                     11538-014"
BASE COUNT       15 a     27 c     24 g      4 t
ORIGIN

  Query Match             71.0%;  Score 14.2;  DB 9;  Length 70;
  Best Local Similarity   84.2%;  Pred. No. 1.2e+04;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 GCTTCCCACCCTGAATGAT 20
           || |||||||| ||||||
Db      37 GCGTCCCACCCAAAATGAT 55


RESULT 5
BH847460/c
LOCUS       BH847460                 96 bp    DNA     linear   GSS 13-JUN-2002
DEFINITION  SALK_054484.25.10.x Arabidopsis thaliana TDNA insertion lines
            Arabidopsis thaliana genomic clone SALK_054484.25.10.x, DNA
            sequence.
ACCESSION   BH847460
VERSION     BH847460.1  GI:21418331
KEYWORDS    GSS.
SOURCE      thale cress.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 96)
  AUTHORS   Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R., Gadrinab
            ,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L., Shinn,P.
            , Zimmerman,J. and Ecker,J.R.
  TITLE     A Sequence-Indexed Library of Insertion Mutations in the
            Arabidopsis Genome
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Joseph R. Ecker
            Salk Institute Genomic Analysis Laboratory (SIGnAL)
            The Salk Institute for Biological Studies
            10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
            Tel: 858 453 4100 x1752
            Fax: 858 558 6379
            Email: ecker@salk.edu
            This is single pass sequence recovered from the left border of
            TDNA. This sequence lies within an annotated intron of At5g43010.
            Class: TDNA tagged.
FEATURES             Location/Qualifiers
     source          1. .96
                     /organism="Arabidopsis thaliana"
                     /strain="Columbia 0"
                     /db_xref="taxon:3702"
                     /clone="SALK_054484.25.10.x"
                     /clone_lib="Arabidopsis thaliana TDNA insertion lines"
                     /note="PCR was performed on Arabidopsis thaliana lines
                     each of which contains  one or more TDNA insertion
                     elements. The resultant fragment for each  line was
                     directly sequenced to determine the genomic sequence at
                     the  site of insertion. Details of the protocols used can
                     be found at  http://signal.salk.edu/tdna_protocols.html"
BASE COUNT       19 a     23 c     21 g     33 t
ORIGIN

  Query Match             71.0%;  Score 14.2;  DB 17;  Length 96;
  Best Local Similarity   84.2%;  Pred. No. 1.3e+04;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 TGCTTCCCACCCTGAATGA 19
           |||||| ||||||| || |
Db      33 TGCTTCACACCCTGGATAA 15


RESULT 6
AZ595851
LOCUS       AZ595851                 41 bp    DNA     linear   GSS 13-DEC-2000
DEFINITION  1M0408B14R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0408B14 R, DNA sequence.
ACCESSION   AZ595851
VERSION     AZ595851.1  GI:11718041
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 41)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0408  row: B  column: 14
            Seq primer: CACACAGGAAACAGCTATGACC
```

```
          Class: plasmid ends
          High quality sequence stop: 41.
FEATURES             Location/Qualifiers
     source          1. .41
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC1M0408B14"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
BASE COUNT        6 a     12 c      8 g     15 t
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 17;  Length 41;
  Best Local Similarity   80.0%;  Pred. No. 1.9e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 TGCTTCCCACCCTGAATGAT 20
           ||||||| || |||| | ||
Db      14 TGCTTCCTACGCTGACTTAT 33


RESULT 7
AZ799423/c
LOCUS       AZ799423                 62 bp    DNA     linear   GSS 16-FEB-2001
DEFINITION  2M0056G22R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC2M0056G22 R, DNA sequence.
ACCESSION   AZ799423
VERSION     AZ799423.1  GI:12950525
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 62)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0056  row: G  column: 22
            Seq primer: CACACAGGAAACAGCTATGACC
```

```
          Class: plasmid ends
          High quality sequence stop: 62.
FEATURES             Location/Qualifiers
     source          1. .62
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC2M0056G22"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
BASE COUNT       11 a     13 c     23 g     15 t
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 17;  Length 62;
  Best Local Similarity   80.0%;  Pred. No. 2.2e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 TGCTTCCCACCCTGAATGAT 20
           |||||||| || |||  ||
Db      29 TGCTTCCCAGCCAGAACAAT 10


RESULT 8
BM565783
LOCUS       BM565783                 68 bp    mRNA    linear   EST 20-FEB-2002
DEFINITION  rt02d09.y1 Pristionchus pacificus mixed stage SL1 TOPO v1 Murphy
            Chiapelli McCarter Pristionchus pacificus cDNA 5', mRNA sequence.
ACCESSION   BM565783
VERSION     BM565783.1  GI:18826253
KEYWORDS    EST.
SOURCE      Pristionchus pacificus.
  ORGANISM  Pristionchus pacificus
            Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
            Neodiplogasteridae; Pristionchus.
REFERENCE   1  (bases 1 to 68)
  AUTHORS   McCarter,J., Clifton,S., Chiapelli,B., Pape,D., Martin,J., Wylie,T.
            , Dante,M., Marra,M., Hillier,L., Kucaba,T., Theising,B., Bowers,Y.
            , Gibbons,M., Ritter,E., Bennett,J., Franklin,C., Tsagareishvili,R.
            , Ronko,I., Kennedy,S., Maguire,L., Beck,C., Underwood,K., Steptoe
            ,M., Allen,M., Person,B., Swaller,T., Harvey,N., Schurk,R., Kohn,S.
            , Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and
            Wilson,R.
  TITLE     The Washington Univ. Nematode EST Project, 1999
  JOURNAL   Unpublished (1999)
COMMENT     Contact: McCarter JP
            The Washington Univ. Nematode EST Project, 1999
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            The library was constructed by Claire Murphy, Brandi Chiapelli, and
            Dr. James McCarter at Washington University, St. Louis. DNA
```

```
            Sequencing by: Washington University Genome Sequencing Center
            Seq primer: -40RP from Gibco.
FEATURES              Location/Qualifiers
     source           1. .68
                      /organism="Pristionchus pacificus"
                      /db_xref="taxon:54126"
                      /clone_lib="Pristionchus pacificus mixed stage SL1 TOPO v1
                      Murphy Chiapelli McCarter"
                      /dev_stage="Mixed stage"
                      /lab_host="DH10B"
                      /note="Vector: pCRII-TOPO (Invitrogen); Site_1: EcoRI;
                      Site_2: EcoRI; The library was constructed by Claire
                      Murphy, Brandi Chiapelli,  and Dr. James McCarter at
                      Washington University, St. Louis. Oligo(dT)-SL1  PCR based
                      library. Pristionchus pacificus mixed stage cDNA PCR
                      products of  size >400 nucleotides containing SL1 on the
                      5' end and oligo(dT) on the 3' end  were non-directionally
                      cloned into pCRII-TOPO(Invitrogen) following the  Topo TA
                      cloning protocol. "
BASE COUNT       15 a     22 c      6 g     25 t
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 13;  Length 68;
  Best Local Similarity   80.0%;  Pred. No. 2.2e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 TGCTTCCCACCCTGAATGAT 20
           ||||||| | |||  |||||
Db       2 TGCTTCCTATCCTCGATGAT 21


RESULT 9
T65253
LOCUS       T65253                    69 bp    mRNA    linear   EST 07-MAR-1995
DEFINITION  yc79b02.r1 Soares infant brain 1NIB Homo sapiens cDNA clone
            IMAGE:21999 5', mRNA sequence.
ACCESSION   T65253
VERSION     T65253.1  GI:674298
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 69)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman
            ,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J.,
            Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston
            ,R., Williamson,A., Wohldmann,P. and Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Insert Size: 2497
            High qality sequence stops: 54 Source: IMAGE Consortium, LLNL This
            clone is available royalty-free through LLNL ; contact the IMAGE
            Consortium (info@image.llnl.gov) for further information.
            Insert Length: 2497   Std Error: 0.00
            Seq primer: M13RP1
            High quality sequence stop: 54.
FEATURES              Location/Qualifiers
     source           1. .69
                      /organism="Homo sapiens"
                      /db_xref="GDB:394346"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:21999"
                      /clone_lib="Soares infant brain 1NIB"
                      /sex="female"
                      /dev_stage="73 days post natal"
                      /lab_host="DH10B (ampicillin resistant)"
                      /note="Organ: whole brain; Vector: Lafmid BA; Site_1: Not
                      I; Site_2: Hind III; 1st strand cDNA was primed with a Not
                      I - oligo(dT) primer [5'
                      AACTGGAAGAATTCGCGGCCGCAGGAATTTTTTTTTTTTTTTT 3']
                      double-stranded cDNA was ligated to Hind III adaptors
                      (Pharmacia), digested with Not I and directionally cloned
                      into the Not I and Hind III sites of the Lafmid BA vector.
                      Library went through one round of normalization. Library
                      constructed by Bento Soares and M.Fatima Bonaldo."
BASE COUNT       19 a     18 c      14 g     18 t
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 14;  Length 69;
  Best Local Similarity   80.0%;  Pred. No. 2.2e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 TGCTTCCCACCCTGAATGAT 20
           |||| |||||||| | || |
Db      34 TGCTCCCCACCCTTACTGCT 53


RESULT 10
BG508809
LOCUS       BG508809                  92 bp    mRNA    linear   EST 28-NOV-2001
DEFINITION  sac90f07.y1 Gm-c1073 Glycine max cDNA clone GENOME SYSTEMS CLONE
            ID: Gm-c1073-278 5', mRNA sequence.
ACCESSION   BG508809
VERSION     BG508809.1  GI:13479466
KEYWORDS    EST.
SOURCE      soybean.
  ORGANISM  Glycine max
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
            Glycine.
REFERENCE   1  (bases 1 to 92)
  AUTHORS   Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Coryell,V., Khanna
            ,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
            Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers
            ,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk
            ,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann
            ,R., Waterston,R. and Wilson,R.
  TITLE     Public Soybean EST Project
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Shoemaker R/Public Soybean EST Project
            Public Soybean EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Putative full length read
            vector to vector length is 93 This clone is available through:
            ResGen, Invitrogen Corp. 2130 South Memorial Parkway Huntsville, AL
            35801 For further information call: (800)-533-4363 or contact via
            email: ccu@resgen.com.
FEATURES              Location/Qualifiers
     source           1. .92
                      /organism="Glycine max"
                      /db_xref="taxon:3847"
                      /clone="GENOME SYSTEMS CLONE ID: Gm-c1073-278"
                      /clone_lib="Gm-c1073"
                      /tissue_type="seedlings induced for symptoms of SDS
                      (Sudden Death Syndrome) disease"
                      /dev_stage="2-3 weeks old"
                      /lab_host="DH10B"
                      /note="Vector: pBluescript II SK+; Site_1: EcoRI; Site_2:
                      XhoI; The cDNA library was constructed from mRNA isolated
                      from 2-3 week old seedlings that were induced for symptoms
                      of SDS (Sudden Death Syndrome) disease by the
                      translocation of culture filtrate of Fusarium solani f.
                      sp. glycines (Plant Cell Report 18:375-380). Cultivar
                      Williams 82 is  susceptible to the disease SDS. Plant
```

tissue (expanded leaves, folded leaves, and new shoots) were collected at 1, 6, 24, and 48 hrs. after inoculation and their mRNA pooled equally for cDNA construction. The library was prepared using the Stratagene pBluescript II SK(+) library construction kit. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dT) sequence with an XhoI restriction site. EcoRI adaptors were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA insert is protected from XhoI digestion via methylation during first strand synthesis. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into E.coli ElectroMax DH10B host cells. Plants were inoculated by Shuxian Li (Glen Hartman lab, University of Illinois). Library was constructed by Reena Philip and Steve Clough (Lila Vodkin lab, University of Illinois)."

BASE COUNT 17 a 17 c 18 g 40 t
ORIGIN

Query Match 68.0%; Score 13.6; DB 12; Length 92;
Best Local Similarity 80.0%; Pred. No. 2.4e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
Qy       1 TGCTTCCCACCCTGAATGAT 20
           |||||||||| | |||| |
Db      60 TGCTTCCCACAATTAATGTT 79
```

RESULT 11
AA885649/c
LOCUS AA885649 97 bp mRNA linear EST 09-JUN-1998
DEFINITION oj32c12.s1 NCI_CGAP_Lu5 Homo sapiens cDNA clone IMAGE:1500022 3' similar to TR:Q13231 Q13231 CHITOTRIOSIDASE PRECURSOR. ;, mRNA sequence.
ACCESSION AA885649
VERSION AA885649.1 GI:3000757
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 97)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html

Trace considered overall poor quality
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html

Trace considered overall poor quality
Insert Length: 671 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 1.
FEATURES Location/Qualifiers
source 1. .97
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1500022"
/clone_lib="NCI_CGAP_Lu5"
/tissue_type="carcinoid"
/lab_host="DH10B"
/note="Organ: lung; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; 1st strand cDNA was prepared from neuroendocrine lung carcinoid, and was then primed with a Not I - oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library is normalized. Library was constructed by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT 22 a 21 c 34 g 20 t
ORIGIN

Query Match 68.0%; Score 13.6; DB 9; Length 97;
Best Local Similarity 80.0%; Pred. No. 2.4e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
Qy       1 TGCTTCCCACCCTGAATGAT 20
           || | |||||||||||  ||
Db      30 TGATCCCCACCCTGAAGTAT 11
```

RESULT 12
AZ478351/c
LOCUS AZ478351 63 bp DNA linear GSS 04-OCT-2000
DEFINITION 1M0298B23F Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC1M0298B23 F, DNA sequence.
ACCESSION AZ478351
VERSION AZ478351.1 GI:10637117
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 63)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D.,Weiss,R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0298 row: B column: 23
Seq primer: CGTTGTAAAACGACGGCCAGT
Class: plasmid ends
High quality sequence stop: 63.
FEATURES Location/Qualifiers
source 1. .63
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0298B23"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA

was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptored DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptored mouse DNA was annealed to adaptored vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

```
BASE COUNT       14 a     13 c     18 g     18 t
ORIGIN

  Query Match            67.0%;  Score 13.4;  DB 17;  Length 63;
  Best Local Similarity   93.3%;  Pred. No. 2.7e+04;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 TGCTTCCCACCCTGA 15
          ||||||||||| |||
Db     49 TGCTTCCCACCATGA 35


RESULT 13
AV970203
LOCUS       AV970203                 70 bp    mRNA    linear   EST 14-MAR-2002
DEFINITION  AV970203 Nori Satoh unpublished cDNA library, cleavage stage embryo
            Ciona intestinalis cDNA clone cicl17p05 5', mRNA sequence.
ACCESSION   AV970203
VERSION     AV970203.1  GI:19459967
KEYWORDS    EST.
SOURCE      Ciona intestinalis.
  ORGANISM  Ciona intestinalis
            Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
            Phlebobranchia; Cionidae; Ciona.
REFERENCE   1  (bases 1 to 70)
  AUTHORS   Satoh,N., Satou,Y., Kohara,Y. and Shin-i,T.
  TITLE     Expressed genes in Ciona intestinalis
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Nori Satoh
            Department of Zoology
            Kyoto University
            Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
            Tel: 81-75-753-4081
            Fax: 81-75-705-1113
            Email: satoh@ascidian.zool.kyoto-u.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .70
                     /organism="Ciona intestinalis"
                     /db_xref="taxon:7719"
                     /clone="cicl17p05"
                     /clone_lib="Nori Satoh unpublished cDNA library, cleavage
                     stage embryo"
                     /tissue_type="whole animal"
                     /dev_stage="cleavage stage embryo"
                     /note="Vector: pBluescript SK"
BASE COUNT       25 a     20 c     12 g     12 t      1 others
ORIGIN

  Query Match            67.0%;  Score 13.4;  DB 10;  Length 70;
  Best Local Similarity   93.3%;  Pred. No. 2.8e+04;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      5 TCCCACCCTGAATGA 19
          ||||||||| |||||
Db     26 TCCCACCCTCAATGA 40


RESULT 14
AZ665595/c
LOCUS       AZ665595                 86 bp    DNA     linear   GSS 14-DEC-2000
DEFINITION  1M0547E07F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0547E07 F, DNA sequence.
ACCESSION   AZ665595
VERSION     AZ665595.1  GI:11802741
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 86)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0547  row: E  column: 07
            Seq primer: CGTTGTAAAACGACGGCCAGT
            Class: plasmid ends
            High quality sequence stop: 86.
FEATURES             Location/Qualifiers
     source          1. .86
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC1M0547E07"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
BASE COUNT       23 a     10 c     28 g     25 t
ORIGIN

  Query Match            67.0%;  Score 13.4;  DB 17;  Length 86;
  Best Local Similarity   93.3%;  Pred. No. 2.9e+04;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 TGCTTCCCACCCTGA 15
          ||||||||||| |||
Db     72 TGCTTCCCACCATGA 58


RESULT 15
```

```
AU076672/c
LOCUS       AU076672                  87 bp    mRNA    linear   EST 04-MAY-2000
DEFINITION  AU076672 Sugano cDNA library Homo sapiens cDNA clone HEP02824
            similar to 5'-end region of Human interferon-gamma receptor mRNA,
            mRNA sequence.
ACCESSION   AU076672
VERSION     AU076672.1  GI:7439153
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 87)
  AUTHORS   Suzuki,Y., Ishihara,D., Sasaki,M., Nakagawa,H., Hata,H., Tsunoda,T.
            , Watanabe,M., Komatsu,T., Ota,T., Isogai,T., Suyama,A. and Sugano
            ,S.
  TITLE     Statistical analysis of the 5' untranslated region of human mRNA
            using 'Oligo-Capped' cDNA libraries
  JOURNAL   Genomics 64 (3), 286-297 (2000)
  MEDLINE   20221373
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp
            Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano
            ,S. Construction and characterization of a full length-enriched and
            a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997)
             This clone was obtained from a 'full length-enriched' cDNA library
            constructed by 'Oligo-Capping' method. The coding region starts
            from the 50 bp upstream to the 3'-end.
FEATURES             Location/Qualifiers
     source          1. .87
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="HEP02824"
                     /clone_lib="Sugano cDNA library"
                     /note="The cDNA was prepared using the anchor primer,
                     H-T11G, from Genhunter"
BASE COUNT       11 a     27 c     28 g     21 t
ORIGIN

  Query Match             67.0%;  Score 13.4;  DB 9;  Length 87;
  Best Local Similarity   93.3%;  Pred. No. 3e+04;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       5 TCCCACCCTGAATGA 19
           || ||||||||||||
Db      80 TCACACCCTGAATGA 66


Search completed: November 26, 2002, 04:08:36
Job time : 766.8 secs
```

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 3, 2002, 14:46:40 ; Search time 302.2 Seconds
(without alignments)
1926.063 Million cell updates/sec

Title: US-09-296-264-15
Perfect score: 20
Sequence: 1 tgggaatagatgaagttgcc 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 995600

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	15.4	77.0	97	1	AF455568	AF455568 Nostoc sp
c	2	13.8	69.0	48	6	I04807	I04807 Sequence 17
	3	13.6	68.0	66	1	AF254834	AF254834 Pyrococcu
c	4	13.4	67.0	30	6	AX188443	AX188443 Sequence
c	5	13.4	67.0	30	6	AX353551	AX353551 Sequence
c	6	13.4	67.0	71	8	GZGAOA1	U51093 Gibberella
	7	13.4	67.0	75	10	MMU77076	U77076 Mus musculu
c	8	13.2	66.0	21	6	A56997	A56997 Sequence 55
c	9	13.2	66.0	43	6	AX483518	AX483518 Sequence
c	10	13.2	66.0	75	8	MISC40	X00891 Yeast mitoc
c	11	13.2	66.0	75	8	MISCTRNK	X04561 Yeast mitoc
c	12	13.2	66.0	75	8	YSCMTTRK2	M27723 S.cerevisia
c	13	13.2	66.0	75	8	YSTMTTRK	K03322 Saccharomyc
	14	13	65.0	51	6	AX189846	AX189846 Sequence
	15	13	65.0	51	6	AX189847	AX189847 Sequence
c	16	12.8	64.0	17	6	AX265839	AX265839 Sequence
	17	12.8	64.0	17	6	AX265840	AX265840 Sequence
	18	12.8	64.0	20	6	AR100287	AR100287 Sequence
	19	12.8	64.0	20	6	AR149942	AR149942 Sequence
	20	12.8	64.0	20	6	AR150230	AR150230 Sequence
c	21	12.8	64.0	31	6	AX248591	AX248591 Sequence
	22	12.8	64.0	41	6	AX214505	AX214505 Sequence
c	23	12.8	64.0	65	6	AX484137	AX484137 Sequence
c	24	12.8	64.0	93	6	AX047817	AX047817 Sequence
c	25	12.8	64.0	93	17	HSMC35E11	X88121 H.sapiens D
	26	12.8	64.0	95	6	A01349	A01349 Human immun
	27	12.8	64.0	95	6	A13130	A13130 Nucleotide
	28	12.8	64.0	95	6	A16103	A16103 oligonucleo
	29	12.6	63.0	23	6	AX430781	AX430781 Sequence
	30	12.6	63.0	23	6	AX486809	AX486809 Sequence
	31	12.6	63.0	33	6	E10044	E10044 Synthetic D
	32	12.6	63.0	33	6	I32970	I32970 Sequence 24
	33	12.6	63.0	33	6	I72604	I72604 Sequence 11
c	34	12.6	63.0	60	6	AR054506	AR054506 Sequence
c	35	12.6	63.0	73	6	AR200235	AR200235 Sequence
c	36	12.6	63.0	79	3	SUPHISAAD	M25333 Sea urchin
c	37	12.6	63.0	87	6	AR007061	AR007061 Sequence
c	38	12.6	63.0	87	6	AR111085	AR111085 Sequence
c	39	12.6	63.0	87	6	I75001	I75001 Sequence 20
	40	12.4	62.0	24	6	AR167462	AR167462 Sequence
	41	12.4	62.0	24	6	AR170511	AR170511 Sequence
	42	12.4	62.0	24	6	AR182985	AR182985 Sequence
	43	12.4	62.0	34	6	AX207565	AX207565 Sequence
	44	12.4	62.0	51	6	AX163410	AX163410 Sequence
c	45	12.4	62.0	55	6	AX483749	AX483749 Sequence

ALIGNMENTS

RESULT 1
AF455568
LOCUS AF455568 97 bp DNA linear BCT 17-JUL-2002
DEFINITION Nostoc sp. HCC 1075 hupSL intergenic region, complete sequence.
ACCESSION AF455568
VERSION AF455568.1 GI:21886668
KEYWORDS .
SOURCE Nostoc sp. HCC 1075.
ORGANISM Nostoc sp. HCC 1075
Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
REFERENCE 1 (bases 1 to 97)
AUTHORS Tamagnini,P., Leitao,E., Almeida,L., Lindberg,P. and Lindblad,P.
TITLE Repetitive sequences within intergenic regions of cyanobacterial uptake hydrogenase genes (hupSL)
JOURNAL Unpublished

```
REFERENCE   2  (bases 1 to 97)
  AUTHORS   Tamagnini,P., Leitao,E., Almeida,L., Lindberg,P. and Lindblad,P.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-DEC-2001) Botany / IBMC, Univ. of Porto, R. do Campo
            Alegre, 823, Porto 4150-180, Portugal
FEATURES             Location/Qualifiers
     source          1. .97
                     /organism="Nostoc sp. HCC 1075"
                     /strain="HCC 1075"
                     /isolate="Mitsui 91911"
                     /db_xref="taxon:197317"
     misc_feature    1. .97
                     /note="hupSL intergenic region"
BASE COUNT       31 a     24 c     24 g     18 t
ORIGIN

  Query Match             77.0%;  Score 15.4;  DB 1;  Length 97;
  Best Local Similarity   94.1%;  Pred. No. 3.7e+03;
  Matches   16;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      2 GGGAATAGATGAAGTTG 18
          ||||||||| ||||||||
Db     21 GGGAATAGGTGAAGTTG 37


RESULT 2
I04807/c
LOCUS       I04807                    48 bp    DNA     linear   PAT 02-DEC-1994
DEFINITION  Sequence 17 from Patent EP 0206783.
ACCESSION   I04807
VERSION     I04807.1  GI:591459
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 48)
  AUTHORS   Thill,G.P., Harpold,M.M. and Tschopp,J.F.
  TITLE     Expression and secretion of polypeptides from saccharomyces
            cerevisiae
  JOURNAL   Patent: EP 0206783-A2 17 30-DEC-1986;
FEATURES             Location/Qualifiers
     source          1. .48
                     /organism="unknown"
BASE COUNT       18 a     13 c      9 g      8 t
ORIGIN

  Query Match             69.0%;  Score 13.8;  DB 6;  Length 48;
  Best Local Similarity   88.2%;  Pred. No. 2.4e+04;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 TGGGAATAGATGAAGTT 17
          ||||| | |||||||||
Db     47 TGGGACTTGATGAAGTT 31


RESULT 3
AF254834
LOCUS       AF254834                  66 bp    mRNA    linear   BCT 21-SEP-2000
DEFINITION  Pyrococcus abyssi box C/D small nucleolar RNA sR2.
ACCESSION   AF254834
VERSION     AF254834.1  GI:10242150
KEYWORDS    .
SOURCE      Pyrococcus abyssi.
  ORGANISM  Pyrococcus abyssi
            Archaea; Euryarchaeota; Thermococci; Thermococcales;
            Thermococcaceae; Pyrococcus.
REFERENCE   1  (bases 1 to 66)
  AUTHORS   Gaspin,C., Cavaille,J., Erauso,G. and Bachellerie,J.P.
  TITLE     Archaeal homologs of eukaryotic methylation guide small nucleolar
            RNAs: lessons from the Pyrococcus genomes
  JOURNAL   J. Mol. Biol. 297 (4), 895-906 (2000)
  MEDLINE   20202371
   PUBMED   10736225

REFERENCE   2  (bases 1 to 66)
  AUTHORS   Gaspin,C., Cavaille,J. and Bachellerie,J.P.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-APR-2000) Unite de Biometrie et Intelligence
            Artificielle, Institut National de la Recherche Agronomique, Chemin
            de Borde-Rouge Auzeville BP 27, Castanet-Tolosan 31326, France
FEATURES             Location/Qualifiers
     source          1. .66
                     /organism="Pyrococcus abyssi"
                     /db_xref="taxon:29292"
                     /map="between PAB0758 and PAB0759"
     snRNA           1. .66
                     /product="box C/D small nucleolar RNA sR2"
                     /note="guides 2'-O methylation on small subunit ribosomal
                     RNA at C766 and G838 inferred from similar snoRNAs in
                     eukaryotes"
BASE COUNT       13 a     10 c     23 g     20 t
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 1;  Length 66;
  Best Local Similarity   80.0%;  Pred. No. 3.1e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 TGGGAATAGATGAAGTTGCC 20
          ||||  ||| ||| ||||||
Db     46 TGGGCTTAGCTGATGTTGCC 65


RESULT 4
AX188443/c
LOCUS       AX188443                  30 bp    DNA     linear   PAT 08-AUG-2001
DEFINITION  Sequence 62 from Patent WO0147954.
ACCESSION   AX188443
VERSION     AX188443.1  GI:15142114
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 30)
  AUTHORS   van Roy,F., Vanlandschoot,A. and Janssens,B.
  TITLE     Novel cdnas encoding catenin-binding proteins with function in
            signalling and/or gene regulation
  JOURNAL   Patent: WO 0147954-A 62 05-JUL-2001;
            Vlaams Interuniversitair Instituut voor Biotechnologie vzw. (BE)
FEATURES             Location/Qualifiers
     source          1. .30
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="primer FVR137F"
BASE COUNT        6 a     13 c      5 g      6 t
ORIGIN

  Query Match             67.0%;  Score 13.4;  DB 6;  Length 30;
  Best Local Similarity   93.3%;  Pred. No. 3.9e+04;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      6 ATAGATGAAGTTGCC 20
          |||| ||||||||||
Db     25 ATAGGTGAAGTTGCC 11


RESULT 5
AX353551/c
LOCUS       AX353551                  30 bp    DNA     linear   PAT 06-FEB-2002
DEFINITION  Sequence 83 from Patent WO0204636.
ACCESSION   AX353551
VERSION     AX353551.1  GI:18618626
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   van Roy,F., Goossens,S., Janssens,B. and Vanpoucke,G.
```

```
  TITLE     Novel _g(a) expressed in heart and testis
  JOURNAL   Patent: WO 0204636-A 83 17-JAN-2002;
            Vlaams Interuniversitair Instituut voor Biotechnologie vzw. (BE)
FEATURES             Location/Qualifiers
     source          1. .30
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="primer MCB137"
BASE COUNT        6 a      13 c       5 g       6 t
ORIGIN

  Query Match             67.0%;  Score 13.4;  DB 6;  Length 30;
  Best Local Similarity   93.3%;  Pred. No. 3.9e+04;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      6 ATAGATGAAGTTGCC 20
          |||| ||||||||||
Db     25 ATAGGTGAAGTTGCC 11


RESULT 6
GZGAOA1/c
LOCUS       GZGAOA1                   71 bp    DNA     linear   PLN 02-JAN-1998
DEFINITION  Gibberella zeae galactose oxidase (gaoA) gene, partial 5' flanking
            region.
ACCESSION   U51093
VERSION     U51093.1  GI:2739397
KEYWORDS    .
SEGMENT     1 of 2
SOURCE      Gibberella zeae strain=DSM 4527.
  ORGANISM  Gibberella zeae
            Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
            Hypocreales; Nectriaceae; Gibberella.
REFERENCE   1  (bases 1 to 71)
  AUTHORS   Niessen,M.L. and Vogel,R.F.
  TITLE     Specific identification of Fusarium graminearum by PCR with gaoA
            targeted primers
  JOURNAL   Syst. Appl. Microbiol. 20, 111-123 (1997)
REFERENCE   2  (bases 1 to 71)
  AUTHORS   Niessen,M.L. and Vogel,R.F.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-MAR-1996) Martin L. Niessen, Lehrstuhl fur Technische
            Mikrobiologie, TU Munchen-Weihenstephan, Freising 85350, Germany
FEATURES             Location/Qualifiers
     source          1. .71
                     /organism="Gibberella zeae"
                     /strain="DSM 4527"
                     /db_xref="taxon:5518"
BASE COUNT       21 a      17 c       8 g      25 t
ORIGIN

  Query Match             67.0%;  Score 13.4;  DB 8;  Length 71;
  Best Local Similarity   93.3%;  Pred. No. 4e+04;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      4 GAATAGATGAAGTTG 18
          ||||||||||||  ||
Db     56 GAATAGATGAAGGTG 42


RESULT 7
MMU77076
LOCUS       MMU77076                  75 bp    DNA     linear   ROD 02-APR-1997
DEFINITION  Mus musculus focal adhesion kinase isoform gene, exon E7aa and
            partial cds.
ACCESSION   U77076
VERSION     U77076.1  GI:1916738
KEYWORDS    .
SOURCE      Mus musculus.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 75)
  AUTHORS   Asano,H., Komiyama,H.K. and Grant,S.G.
  TITLE     Isoforms of FAK in brain
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 75)
  AUTHORS   Asano,H., Komiyama,H.K. and Grant,S.G.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-NOV-1996) CGR, Edinburgh University, West Mains Road,
            Edinburgh, Scotland EH9 3JQ, U.K
FEATURES             Location/Qualifiers
     source          1. .75
                     /organism="Mus musculus"
                     /strain="129SJVMus"
                     /db_xref="taxon:10090"
     intron          1. .29
     CDS             <31. .>51
                     /note="FAK"
                     /codon_start=3
                     /product="focal adhesion kinase isoform"
                     /protein_id="AAB51231.1"
                     /db_xref="GI:1916739"
                     /translation="SYGIDE"
     exon            31. .51
                     /note="exon E7aa"
     intron          52. .75
BASE COUNT       24 a      11 c      13 g      27 t
ORIGIN

  Query Match             67.0%;  Score 13.4;  DB 10;  Length 75;
  Best Local Similarity   93.3%;  Pred. No. 4e+04;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      3 GGAATAGATGAAGTT 17
          |||||||||||||  |
Db     39 GGAATAGATGAAGGT 53


RESULT 8
A56997/c
LOCUS       A56997                    21 bp    DNA     linear   PAT 03-MAR-1998
DEFINITION  Sequence 55 from Patent WO9629091.
ACCESSION   A56997
VERSION     A56997.1  GI:3712980
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 21)
  AUTHORS   Stanley,M.A. and Scarpini,C.G.
  TITLE     TREATMENT OF PAPILLOMAVIRUS-ASSOCIATED LESIONS USING INTERLEUKIN-12
  JOURNAL   Patent: WO 9629091-A 55 26-SEP-1996;
            UNIV CAMBRIDGE TECH (GB)
COMMENT     Other publication AU 5151596 961008.
FEATURES             Location/Qualifiers
     source          1. .21
                     /organism="unidentified"
                     /db_xref="taxon:32644"
BASE COUNT        3 a       8 c       3 g       7 t
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 6;  Length 21;
  Best Local Similarity   83.3%;  Pred. No. 4.8e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 TGGGAATAGATGAAGTTG 18
          ||||| |||| ||||| |
Db     18 TGGGAGTAGACGAAGTAG 1


RESULT 9
AX483518/c
LOCUS       AX483518                  43 bp    DNA     linear   PAT 16-AUG-2002
DEFINITION  Sequence 818 from Patent WO02053728.
ACCESSION   AX483518
```

```
VERSION     AX483518.1  GI:22317938
KEYWORDS    .
SOURCE      Candida albicans.
  ORGANISM  Candida albicans
            Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
            Saccharomycetales; mitosporic Saccharomycetales; Candida.
REFERENCE   1
  AUTHORS   Roemer,T., Jiang,B., Boone,C., Bussey,H. and Ohlsen,K.L.
  TITLE     Gene disruption methodologies for drug target discovery
  JOURNAL   Patent: WO 02053728-A 818 11-JUL-2002;
            Elitra Pharmaceuticals, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .43
                     /organism="Candida albicans"
                     /db_xref="taxon:5476"
BASE COUNT       19 a      10 c       0 g      14 t
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 6;  Length 43;
  Best Local Similarity   83.3%;  Pred. No. 5e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 TGGGAATAGATGAAGTTG 18
           | ||||||||||  |||
Db      28 TTGGAATAGATGAGTTTG 11


RESULT 10
MISC40/c
LOCUS       MISC40                    75 bp    mRNA    linear   PLN 02-AUG-1994
DEFINITION  Yeast mitochondrial transfer RNA-Lys sequence.
ACCESSION   X00891
VERSION     X00891.1  GI:13560
KEYWORDS    transfer RNA; transfer RNA-Lys.
SOURCE      Saccharomyces cerevisiae.
  ORGANISM  Mitochondrion Saccharomyces cerevisiae
            Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
            Saccharomycetales; Saccharomycetaceae; Saccharomyces.
REFERENCE   1  (bases 1 to 75)
  AUTHORS   Martin,R.P., Sibler,A.P., Bordonne,R. and Dirheimer,G.
  TITLE     Yeast mitochondrial tRNAs: structure, coding properties, and gene
            organization
  JOURNAL   Mol. Biol. (Mosk.) 17, 915-936 (1983)
FEATURES             Location/Qualifiers
     source          1. .75
                     /organism="Saccharomyces cerevisiae"
                     /organelle="mitochondrion"
                     /db_xref="taxon:4932"
     tRNA            1. .75
                     /product="tRNA-Lys"
     modified_base   16
                     /note="dihydrouridine"
                     /mod_base=d
     modified_base   19
                     /note="dihydrouridine"
                     /mod_base=d
     modified_base   27
                     /note="pseudouridine"
                     /mod_base=p
     modified_base   30
                     /note="pseudouridine"
                     /mod_base=p
     modified_base   36
                     /note="N-((9-beta-D-ribofuranosyl- purine-6-yl)carbamoyl
                     threonine"
                     /mod_base=OTHER
     modified_base   53
                     /note="5-methyluridine"
                     /mod_base=t
     modified_base   54
                     /note="pseudouridine"
                     /mod_base=p
BASE COUNT       22 a      16 c      12 g      25 t
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 8;  Length 75;
  Best Local Similarity   83.3%;  Pred. No. 5.1e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 TGGGAATAGATGAAGTTG 18
           || ||||||  || |||||
Db      72 TGAGAATAGCTGGAGTTG 55


RESULT 11
MISCTRNK/c
LOCUS       MISCTRNK                  75 bp    mRNA    linear   PLN 02-AUG-1994
DEFINITION  Yeast mitochondrial transfer RNA-Lys(u*uu).
ACCESSION   X04561
VERSION     X04561.1  GI:13601
KEYWORDS    transfer RNA; transfer RNA-Lys.
SOURCE      Saccharomyces cerevisiae.
  ORGANISM  Mitochondrion Saccharomyces cerevisiae
            Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
            Saccharomycetales; Saccharomycetaceae; Saccharomyces.
REFERENCE   1  (bases 1 to 75)
  AUTHORS   Martin,R.P., Sibler,A.P., Bordonne,R. and Dirheimer,G.
  TITLE     Yeast mitochondrial tRNAs: structure, coding properties, and gene
            organization
  JOURNAL   Mol. Biol. (Mosk.) 17, 915-936 (1983)
FEATURES             Location/Qualifiers
     source          1. .75
                     /organism="Saccharomyces cerevisiae"
                     /organelle="mitochondrion"
                     /db_xref="taxon:4932"
     modified_base   16
                     /note="dihydrouridine"
                     /mod_base=d
     modified_base   19
                     /note="dihydrouridine"
                     /mod_base=d
     modified_base   27
                     /note="pseudouridine"
                     /mod_base=p
     modified_base   30
                     /note="pseudouridine"
                     /mod_base=p
     misc_feature    33. .35
                     /note="u*uu anticodon"
     modified_base   33
                     /note="unknown modification"
                     /mod_base=OTHER
     modified_base   36
                     /note="N-((9-beta-D-ribofuranosylpurine-6-yl)
                     carbamoyl)threonine"
                     /mod_base=OTHER
     modified_base   53
                     /note="5-methyluridine"
                     /mod_base=t
     modified_base   54
                     /note="pseudouridine"
                     /mod_base=p
BASE COUNT       22 a      16 c      12 g      25 t
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 8;  Length 75;
  Best Local Similarity   83.3%;  Pred. No. 5.1e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 TGGGAATAGATGAAGTTG 18
           || ||||||  || |||||
Db      72 TGAGAATAGCTGGAGTTG 55


RESULT 12
YSCMTTRK2/c
```

```
LOCUS       YSCMTTRK2                 75 bp    tRNA    linear   PLN 18-APR-1994
DEFINITION  S.cerevisiae mitochondrial Lys-2-tRNA.
ACCESSION   M27723
VERSION     M27723.1  GI:176382
KEYWORDS    transfer RNA; transfer RNA-Lys-II.
SOURCE      S.cerevisiae mitochondrial tRNA.
  ORGANISM  Mitochondrion Saccharomyces cerevisiae
            Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
            Saccharomycetales; Saccharomycetaceae; Saccharomyces.
REFERENCE   1  (bases 1 to 75)
  AUTHORS   Sibler,A.P., Dirheimer,G. and Martin,R.P.
  TITLE     Codon reading patterns in Saccharomyces cerevisiae mitochondria
            based on sequences of mitochondrial tRNAs
  JOURNAL   FEBS Lett. 194 (1), 131-138 (1986)
  MEDLINE   86082329
   PUBMED   2416594
FEATURES             Location/Qualifiers
     source          1. .75
                     /organism="Saccharomyces cerevisiae"
                     /organelle="mitochondrion"
                     /db_xref="taxon:4932"
     tRNA            1. .75
                     /product="tRNA-Lys"
                     /note="codon recognized: AAA; Lys-2-tRNA"
     modified_base   16
                     /mod_base=d
     modified_base   19
                     /mod_base=d
     modified_base   27
                     /mod_base=p
     modified_base   30
                     /mod_base=p
     modified_base   36
                     /mod_base=t6a
     modified_base   54
                     /mod_base=p
BASE COUNT       22 a     16 c     12 g     23 t      2 others
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 8;  Length 75;
  Best Local Similarity   83.3%;  Pred. No. 5.1e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 TGGGAATAGATGAAGTTG 18
          || |||||| || |||||
Db     72 TGAGAATAGCTGGAGTTG 55


RESULT 13
YSTMTTRK/c
LOCUS       YSTMTTRK                  75 bp    tRNA    linear   PLN 13-MAY-1994
DEFINITION  Saccharomyces cerevisiae mitochondrial Lys-tRNA.
ACCESSION   K03322
VERSION     K03322.1  GI:176475
KEYWORDS    transfer RNA; transfer RNA-Lys.
SOURCE      Yeast tRNA.
  ORGANISM  Mitochondrion Saccharomyces cerevisiae
            Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
            Saccharomycetales; Saccharomycetaceae; Saccharomyces.
REFERENCE   1  (bases 1 to 75)
  AUTHORS   Martin,R.P., Sibler,A.-P., Bordonne,R. and Dirheimer,G.
  TITLE     Yeast mitochondrial tRNAs: structure, coding properties, and gene
            organization
  JOURNAL   Mol. Biol. 17, 915-936 (1983)
FEATURES             Location/Qualifiers
     source          1. .75
                     /organism="Saccharomyces cerevisiae"
                     /organelle="mitochondrion"
                     /db_xref="taxon:4932"
     tRNA            1. .75
                     /product="tRNA-Lys"
                     /anticodon=(pos:33. .35,aa:Lys)
     modified_base   16
                     /mod_base=d
     modified_base   19
                     /mod_base=d
     modified_base   27
                     /mod_base=p
     modified_base   30
                     /mod_base=p
     modified_base   36
                     /mod_base=t6a
     modified_base   54
                     /mod_base=p
BASE COUNT       22 a     16 c     12 g     25 t
ORIGIN      5' end of mature Lys-tRNA.

  Query Match             66.0%;  Score 13.2;  DB 8;  Length 75;
  Best Local Similarity   83.3%;  Pred. No. 5.1e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 TGGGAATAGATGAAGTTG 18
          || |||||| || |||||
Db     72 TGAGAATAGCTGGAGTTG 55


RESULT 14
AX189846
LOCUS       AX189846                  51 bp    DNA     linear   PAT 08-AUG-2001
DEFINITION  Sequence 25 from Patent WO0147942.
ACCESSION   AX189846
VERSION     AX189846.1  GI:15143217
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 51)
  AUTHORS   Shimkets,R.A. and Leach,M.
  TITLE     Nucleic acids containing single nucleotide polymorphisms and
            methods of use thereof
  JOURNAL   Patent: WO 0147942-A 25 05-JUL-2001;
            Curagen Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .51
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /note="1 of 2 allelic variants (26 is other
                     entry)-Accession number cg43114113"
BASE COUNT       16 a     15 c     14 g      6 t
ORIGIN

  Query Match             65.0%;  Score 13;  DB 6;  Length 51;
  Best Local Similarity   100.0%;  Pred. No. 6.3e+04;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      8 AGATGAAGTTGCC 20
          |||||||||||||
Db      1 AGATGAAGTTGCC 13


RESULT 15
AX189847
LOCUS       AX189847                  51 bp    DNA     linear   PAT 08-AUG-2001
DEFINITION  Sequence 26 from Patent WO0147942.
ACCESSION   AX189847
VERSION     AX189847.1  GI:15143218
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 51)
  AUTHORS   Shimkets,R.A. and Leach,M.
  TITLE     Nucleic acids containing single nucleotide polymorphisms and
            methods of use thereof
```

```
  JOURNAL   Patent: WO 0147942-A 26 05-JUL-2001;
            Curagen Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .51
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /note="2 of 2 allelic variants (25 is other
                     entry)~Accession number cg43114113"
BASE COUNT       17 a     15 c     13 g      6 t
ORIGIN

  Query Match             65.0%;  Score 13;  DB 6;  Length 51;
  Best Local Similarity   100.0%;  Pred. No. 6.3e+04;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       8 AGATGAAGTTGCC 20
           |||||||||||||
Db       1 AGATGAAGTTGCC 13


Search completed: December  3, 2002, 18:13:38
Job time : 310.2 secs
```

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 05:52:31 ; Search time 98.55 Seconds
(without alignments)
457.027 Million cell updates/sec

Title: US-09-296-264-15
Perfect score: 20
Sequence: 1 tgggaatagatgaagttgcc 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 2390332

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	20	100.0	20	21	AAZ31445	Human neuropilin m
c	2	16.4	82.0	65	24	ABN32026	Rat spliced transc
c	3	14.8	74.0	60	24	ABN50325	Human spliced tran
c	4	14.4	72.0	60	18	AAT72799	Apo-2 ligand oligo
c	5	14.4	72.0	60	20	AAX86990	Human Apo-2 lignd
c	6	14.4	72.0	60	21	AAA07428	Probe for human Ap
	7	14.2	71.0	20	13	AAQ25471	Purine rich CMV ta
c	8	14.2	71.0	20	21	AAZ43835	Human adult skin c
c	9	14.2	71.0	60	24	ABN39186	Human spliced tran
	10	13.8	69.0	26	24	ABN80359	PCR primer 915-F2.
	11	13.8	69.0	26	24	ABL51232	Thermotolerant rib
	12	13.8	69.0	28	24	ABN80358	PCR primer 915-F1.
	13	13.8	69.0	28	24	ABL51231	Thermotolerant rib
	14	13.8	69.0	37	21	AAA62774	Endoglucanase PCR
	15	13.8	69.0	60	24	ABN48099	Human spliced tran
	16	13.8	69.0	89	22	AAL35015	Human musculoskele
	17	13.8	69.0	96	21	AAA93426	Caspase-3 PCR prim
	18	13.8	69.0	96	21	AAA93430	Caspase-3 PCR prim
	19	13.8	69.0	96	21	AAA27644	Caspase-3 biosenso
	20	13.8	69.0	96	21	AAA27648	Caspase-3 biosenso
c	21	13.6	68.0	65	24	ABN29782	Rat spliced transc
	22	13.6	68.0	65	24	ABN52842	Mouse spliced tran
c	23	13.4	67.0	30	22	AAD12667	Human alphaN-caten
c	24	13.4	67.0	30	22	AAC88752	Human catenin-bind
c	25	13.4	67.0	30	24	ABK41550	Human alpha-cateni
c	26	13.4	67.0	78	22	ABA73192	Human foetal liver
c	27	13.4	67.0	78	22	ABA38631	Probe #17097 for g
c	28	13.4	67.0	78	22	AAK21627	Human brain expres
c	29	13.4	67.0	78	22	AAK47792	Human bone marrow
c	30	13.4	67.0	78	22	AAI53621	Probe #22307 used
c	31	13.4	67.0	78	24	ABS21802	Human genome-deriv
c	32	13.2	66.0	21	17	AAT39018	Tumour necrosis fa
c	33	13.2	66.0	31	22	AAI30182	Human single nucle
c	34	13.2	66.0	40	14	AAQ51637	Primer LyHNII40.
c	35	13.2	66.0	60	24	ABN49926	Human spliced tran
c	36	13.2	66.0	65	24	ABN57366	Mouse spliced tran
	37	13	65.0	17	24	ABN06740	Human GDMLP-1 17-m
	38	13	65.0	17	24	ABN06741	Human GDMLP-1 17-m
	39	13	65.0	17	24	ABN06742	Human GDMLP-1 17-m
	40	13	65.0	17	24	ABN06743	Human GDMLP-1 17-m
	41	13	65.0	17	24	ABN06744	Human GDMLP-1 17-m
	42	13	65.0	25	24	ABN11632	Human GDMLP-1 25-m
	43	13	65.0	25	24	ABN11633	Human GDMLP-1 25-m
	44	13	65.0	25	24	ABN11634	Human GDMLP-1 25-m
	45	13	65.0	25	24	ABN11635	Human GDMLP-1 25-m

ALIGNMENTS

RESULT 1
AAZ31445
ID AAZ31445 standard; DNA; 20 BP.
XX
AC AAZ31445;
XX
DT 07-FEB-2000 (first entry)
XX
DE Human neuropilin mRNA specific antisense oligo GTI3615.
XX
KW Neuropilin; human; growth; metastasis; tumor; neovascularisation;
KW cancer; papilloma; diabetic retinopathy; antisense; ss.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9955855-A2.
XX
PD 04-NOV-1999.
XX
PF 23-APR-1999; 99WO-CA00324.
XX
PR 23-APR-1998; 98US-0082791.
XX
PA (GENE-) GENESENSE TECHNOLOGIES INC.
XX
PI Wright JA, Young AH, Lee YS;
XX
DR WPI; 2000-023357/02.
XX
PT Antisense oligonucleotides that inhibit neuropilin expression, useful
PT for treating cancer -

XX
PS Claim 4; Page 16; 57pp; English.
XX
CC Sequences AAZ31431-460 represent antisense oligonucleotides which
CC inhibit human neuropilin expression. The antisense oligonucleotides can
CC be used to inhibit the growth or metastasis of a mammalian tumor and
CC inhibit neovascularisation. The oligonucleotides may be used to treat
CC various forms of cancers or tumors, such as sarcomas, melanomas,
CC adenomas, carcinomas of solid tissue, hypoxic tumors, squamous cell
CC carcinomas of the mouth, throat, larynx and lung, genitourinary cancers
CC such as cervical and bladder cancer, hematopoietic cancers, colon cancer,
CC breast cancer, pancreatic cancer, renal cancer, brain cancer, skin
CC cancer, liver cancer, head and neck cancers, and nervous system cancers,
CC as well as benign lesions such as papillomas. The methods may be used to
CC treat neovascularisation disorders such as diabetic retinopathy, and
CC retinopathy of prematurity and age related macular degeneration.
XX
SQ Sequence 20 BP; 6 A; 2 C; 7 G; 5 T; 0 other;

Query Match 100.0%; Score 20; DB 21; Length 20;
Best Local Similarity 100.0%; Pred. No. 3;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TGGGAATAGATGAAGTTGCC 20
||||||||||||||||||||
Db 1 TGGGAATAGATGAAGTTGCC 20

RESULT 2
ABN32026/c
ID ABN32026 standard; DNA; 65 BP.
XX
AC ABN32026;
XX
DT 15-JUL-2002 (first entry)
XX
DE Rat spliced transcript detection oligonucleotide SEQ ID NO:4774.
XX
KW Human; mouse; rat; splice transcript; detection; RNA transcript;
KW splice variant; transcriptome; oligonucleotide library; ss.
XX
OS Rattus norvegicus.
XX
PN WO200210449-A2.
XX
PD 07-FEB-2002.
XX
PF 20-JUL-2001; 2001WO-IB01903.
XX
PR 28-JUL-2000; 2000US-221607P.
PR 02-MAY-2001; 2001US-287724P.
XX
PA (COMP-) COMPUGEN INC.
XX
PI Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
XX
DR WPI; 2002-257383/30.
XX
PT New oligonucleotide libraries comprising oligonucleotides which
PT selectively hybridize to mRNAs transcribed from a transcription unit of
PT a genome, useful for detecting tissue-, pathology-, and
PT developmental-specific genes -
XX
PS Example 1; SEQ ID 4774; 47pp; English.
XX
CC The present invention describes oligonucleotide libraries for detecting
CC messenger RNAs that populate a (sub-)transcriptome, where the
CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC transcription units that populate a genome. The library comprises
CC several oligonucleotides, each capable of hybridising selectively to a
CC set of messenger RNAs transcribed from a given transcription unit of
CC the genome, which encodes one or more messenger RNA splice variants.
CC The oligonucleotide libraries are useful for detecting mRNAs from a
CC biological sample, in expression profiling studies, in qualitatively or
CC quantitatively characterising the corresponding transcriptome, and in
CC detecting RNA transcripts and splice variants of human or animal
CC transcriptomes. The libraries may also be used as specialised mini
CC libraries to detect transcripts of a sub-transcriptome under a
CC particular biological or pathological state, and so allowing the
CC detection of tissue- and pathology-specific genes such as those genes
CC only expressed in specific tissue under a specific pathological
CC condition; to detect developmental specific genes; and to detect RNA
CC transcripts and splice variants of a transcriptome of a patient suffering
CC from a particular disorder. ABN27253 to ABN59589 represent
CC oligonucleotide sequences from rats, humans and mice, which are used in
CC the exemplification of the present invention.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 65 BP; 19 A; 22 C; 10 G; 14 T; 0 other;

Query Match 82.0%; Score 16.4; DB 24; Length 65;
Best Local Similarity 94.4%; Pred. No. 1.8e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TGGGAATAGATGAAGTTG 18
||||||| ||||||||||
Db 39 TGGGAATGGATGAAGTTG 22

RESULT 3
ABN50325/c
ID ABN50325 standard; DNA; 60 BP.
XX
AC ABN50325;
XX
DT 15-JUL-2002 (first entry)
XX
DE Human spliced transcript detection oligonucleotide SEQ ID NO:23073.
XX
KW Human; mouse; rat; splice transcript; detection; RNA transcript;
KW splice variant; transcriptome; oligonucleotide library; ss.
XX
OS Homo sapiens.
XX
PN WO200210449-A2.
XX
PD 07-FEB-2002.
XX
PF 20-JUL-2001; 2001WO-IB01903.
XX
PR 28-JUL-2000; 2000US-221607P.
PR 02-MAY-2001; 2001US-287724P.
XX
PA (COMP-) COMPUGEN INC.
XX
PI Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
XX
DR WPI; 2002-257383/30.
XX
PT New oligonucleotide libraries comprising oligonucleotides which
PT selectively hybridize to mRNAs transcribed from a transcription unit of
PT a genome, useful for detecting tissue-, pathology-, and
PT developmental-specific genes -
XX
PS Example 1; SEQ ID 23073; 47pp; English.
XX
CC The present invention describes oligonucleotide libraries for detecting
CC messenger RNAs that populate a (sub-)transcriptome, where the
CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC transcription units that populate a genome. The library comprises
CC several oligonucleotides, each capable of hybridising selectively to a
CC set of messenger RNAs transcribed from a given transcription unit of
CC the genome, which encodes one or more messenger RNA splice variants.
CC The oligonucleotide libraries are useful for detecting mRNAs from a

```
CC   biological sample, in expression profiling studies, in qualitatively or
CC   quantitatively characterising the corresponding transcriptome, and in
CC   detecting RNA transcripts and splice variants of human or animal
CC   transcriptomes. The libraries may also be used as specialised mini
CC   libraries to detect transcripts of a sub-transcriptome under a
CC   particular biological or pathological state, and so allowing the
CC   detection of tissue- and pathology-specific genes such as those genes
CC   only expressed in specific tissue under a specific pathological
CC   condition; to detect developmental specific genes; and to detect RNA
CC   transcripts and splice variants of a transcriptome of a patient suffering
CC   from a particular disorder. ABN27253 to ABN59589 represent
CC   oligonucleotide sequences from rats, humans and mice, which are used in
CC   the exemplification of the present invention.
CC   N.B. The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 60 BP; 18 A; 19 C; 3 G; 20 T; 0 other;

  Query Match             74.0%;  Score 14.8;  DB 24;  Length 60;
  Best Local Similarity   88.9%;  Pred. No. 1e+03;
  Matches   16;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 TGGGAATAGATGAAGTTG 18
          |||||||||||| | |||
Db     34 TGGGAATAGATGCACTTG 17


RESULT 4
AAT72799/c
ID   AAT72799 standard; DNA; 60 BP.
XX
AC   AAT72799;
XX
DT   22-SEP-1997  (first entry)
XX
DE   Apo-2 ligand oligonucleotide probe.
XX
KW   Apo-2 ligand; cytokine; apoptosis; breast cancer; colon cancer;
KW   therapy; probe; ss.
XX
OS   Synthetic.
XX
PN   WO9725428-A1.
XX
PD   17-JUL-1997.
XX
PF   08-JAN-1997;   97WO-US00272.
XX
PR   09-JAN-1996;   96US-0584031.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Ashkenazi AJ,  Chuntharapai A,  Kim KJ;
XX
DR   WPI; 1997-372867/34.
XX
PT   Novel cytokine, Apo-2 ligand and corresponding DNA - used to induce
PT   apoptosis for the treatment of breast and colon cancer
XX
PS   Example 1; Page 31; 72pp; English.
XX
CC   An oligonucleotide probe (AAT72799) is based on HHEA47M (AAT72797), an
CC   EST that shows homology to human fas/Apo-1 ligand.  It was used to
CC   rescreen clones that had previously been isolated from a human
CC   placental cDNA library using another HHEA47M-based probe (AAT72798).
CC   A cDNA clone (AAT72796) coding for a novel cytokine, designated Apo-2
CC   ligand (AAW19777), was obtd.
XX
SQ   Sequence 60 BP; 19 A; 13 C; 10 G; 18 T; 0 other;

  Query Match             72.0%;  Score 14.4;  DB 18;  Length 60;
  Best Local Similarity   93.8%;  Pred. No. 1.6e+03;
  Matches   15;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 TGGGAATAGATGAAGT 16
          |||||||||||| |||
Db     47 TGGGAATAGATGTAGT 32


RESULT 5
AAX86990/c
ID   AAX86990 standard; DNA; 60 BP.
XX
AC   AAX86990;
XX
DT   24-SEP-1999  (first entry)
XX
DE   Human Apo-2 lignd cDNA hybridising probe.
XX
KW   Cytokine; Apo-2 ligand; Apo-2L; apoptosis; cancer; autoimmune disorder;
KW   lupus; immune-mediated glomerular nephritis; human; probe; ss.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
PN   WO9936535-A1.
XX
PD   22-JUL-1999.
XX
PF   15-JAN-1999;   99WO-US01039.
XX
PR   15-APR-1998;   98US-0060533.
PR   15-JAN-1998;   98US-0007886.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Ashkenazi AJ,  Kelley RF,  O'Connel MT,  Pitti RM;
PI   Schwall RH;
XX
DR   WPI; 1999-444397/37.
XX
PT   A novel cytokine, designated Apo-2 ligand, useful for inducing
PT   apoptosis in mammalian cancer cells
XX
PS   Example 1; Page 32; 86pp; English.
XX
CC   The invention relates to a novel human cytokine, designated Apo-2 ligand
CC   (Apo-2L). The Apo-2L polypeptide can be produced by standard recombinant
CC   methodology. Apo-2L is useful for inducing apoptosis in mammalian cancer
CC   cells. This is useful for the treatment of cancer. Apo-2L can be used to
CC   induce apoptosis for pathological conditions characterized by decreased
CC   levels of apoptosis, e.g. autoimmune disorders like lupus and immune-
CC   mediated glomerular nephritis and cancer. Apo-2L and its nucleic acid
CC   coding sequence can also be used in quantitative and screening
CC   diagnostic techniques. Anti-Apo-2L antibodies can be used for treating
CC   diseases associated with increased apoptosis. Sequences AAX86989-990
CC   represent probes for the isolation of human Apo-2L cDNA.
XX
SQ   Sequence 60 BP; 19 A; 13 C; 10 G; 18 T; 0 other;

  Query Match             72.0%;  Score 14.4;  DB 20;  Length 60;
  Best Local Similarity   93.8%;  Pred. No. 1.6e+03;
  Matches   15;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 TGGGAATAGATGAAGT 16
          |||||||||||| |||
Db     47 TGGGAATAGATGTAGT 32


RESULT 6
AAA07428/c
ID   AAA07428 standard; DNA; 60 BP.
XX
AC   AAA07428;
XX
```

```
DT   10-JUL-2000  (first entry)
XX
DE   Probe for human Apo-2 ligand protein coding sequence.
XX
KW   Apo-2 ligand; human; monoclonal antibody; hybridoma cell line; diagnosis;
KW   therapy; apoptosis; cancer; probe; ss.
XX
OS   Homo sapiens.
XX
PN   US6046048-A.
XX
PD   04-APR-2000.
XX
PF   08-JAN-1997;   97US-0780496.
XX
PR   09-JAN-1996;   96US-0009755.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Kim KJ,  Ashkenazi AJ,  Chuntharapai A;
XX
DR   WPI; 2000-282690/24.
XX
PT   New isolated monoclonal antibodies having antigen specificity for Apo-2
PT   ligand, e.g. 2G6, 2E11 or 5C2, useful for detecting the expression of
PT   Apo-2 ligand serum, and for treating diseases associated with increased
PT   apoptosis  -
XX
PS   Example 1; Column 31-32; 46pp; English.
XX
CC   This sequence represents a probe used to isolate the human Apo-2 ligand
CC   protein coding sequence. The Apo-2 ligand protein is recognised
CC   by monoclonal antibodies produced by the hybridoma cell lines of the
CC   invention. The hybridoma cell lines are deposited under the American Type
CC   Culture Collection Accession Numbers: ATCC HB-12256, HB-12257, HB-12258
CC   and HB-12259. The Apo-2 ligand antibodies may be used in diagnostic
CC   assays for Apo-2 ligand, e.g. detecting its expression in specific cells,
CC   tissues, or serum. The antibodies may also be employed as therapeutics.
CC   For instance, anti-Apo-2 ligand antibodies which block Apo-2 ligand
CC   activity, like Apo-2 ligand-induced apoptosis, may be employed to treat
CC   pathological conditions or diseases associated with increased
CC   apoptosis. They are also useful for the affinity purification of Apo-2
CC   ligand from recombinant cell culture or natural sources. The Apo-2
CC   ligand itself may be used to treat diseases e.g. cancer, by inducing
CC   apoptosis in cells.
XX
SQ   Sequence 60 BP; 19 A; 13 C; 10 G; 18 T; 0 other;

  Query Match             72.0%;  Score 14.4;  DB 21;  Length 60;
  Best Local Similarity   93.8%;  Pred. No. 1.6e+03;
  Matches   15;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 TGGGAATAGATGAAGT 16
           |||||||||||| |||
Db      47 TGGGAATAGATGTAGT 32


RESULT 7
AAQ25471
ID   AAQ25471 standard; DNA; 20 BP.
XX
AC   AAQ25471;
XX
DT   07-DEC-1992  (first entry)
XX
DE   Purine rich CMV target duplex sequence.
XX
KW   Target; Cytomegalovirus; Herpes simplex virus; AIDS; triplex; HIV;
KW   hepatitis; malignancy; ds.
XX
OS   Synthetic.
XX
PN   WO9209705-A.
XX
PD   11-JUN-1992.
XX
PF   25-NOV-1991;   91WO-US08811.
XX
PR   23-NOV-1990;   90US-0617907.
PR   18-JAN-1991;   91US-0643382.
PR   08-APR-1991;   91US-0683420.
PR   17-APR-1991;   91US-0686544.
PR   17-APR-1991;   91US-0686546.
PR   17-APR-1991;   91US-0686547.
PR   27-SEP-1991;   91US-0766733.
XX
PA   (GILE-) GILEAD SCI INC.
XX
PI   Froehler B,  Krawczyk S,  Matteucci MD,  Milligan J;
XX
DR   WPI; 1992-217083/26.
XX
PT   New oligomers contg. modified bases - which form a triplex with
PT   G-C doublet in a DNA duplex, for treating and diagnosing HIV,
PT   hepatitis, herpes, malignancy and inflammation
XX
PS   Claim 11; Page 63; 77pp; English.
XX
CC   The sequence depicts cytomegalovirus beginning at nucleotide
CC   176.  The sequence is a viral duplex sequence which contains a
CC   purine-rich region concentrated on one chain of the duplex.  The
CC   sequence may be prepd. by standard DNA synthesis.  The CMV duplex
CC   sequence is used as a target for novel oligomers which are capable
CC   of forming a triplex at physiological pH by coupling into the major
CC   groove of the DNA duplex. Five such oligomers CMV801-5 are capable
CC   of forming a triplex with this sequence.  The oligomers are used in
CC   the diagnosis and therapy of CMV infection. Similar oligomers may be
CC   used to target viral DNA duplexes specific for HIV, herpes and
CC   malignancy.  The triple helices form under mild conditions thus assays
CC   may be carried out without subjecting the test specimen to harsh
CC   conditions.  The oligomer is able to inhibit gene expression, as
CC   verified by in vitro systems.
CC   See also AAQ25452-25501 and AAQ30226-448.
XX
SQ   Sequence 20 BP; 7 A; 3 C; 8 G; 2 T; 0 other;

  Query Match             71.0%;  Score 14.2;  DB 13;  Length 20;
  Best Local Similarity   84.2%;  Pred. No. 1.8e+03;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 GGGAATAGATGAAGTTGCC 20
           ||||| ||| |||||| ||
Db       2 GGGAAAAGAGGAAGTTCCC 20


RESULT 8
AAZ43835/c
ID   AAZ43835 standard; DNA; 20 BP.
XX
AC   AAZ43835;
XX
DT   10-MAR-2000  (first entry)
XX
DE   Human adult skin cDNA clone vd4_1 DNA probe.
XX
KW   Human; secreted protein; treatment; nutritional activity; cytokine;
KW   cell proliferation; cell differentiation; hematopoiesis regulation;
KW   tissue growth; activin; inhibin; chemotactic; chemokinetic; hemostatic;
KW   thrombolytic; anti-inflammatory; invasion suppressor; tumor inhibition;
KW   gene therapy; ss.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
PN   WO9955721-A1.
XX
```

PD 04-NOV-1999.
XX
PF 23-APR-1999; 99WO-US08504.
XX
PR 24-APR-1998; 98US-0082904.
PR 11-JUN-1998; 98US-0088994.
PR 12-JUN-1998; 98US-0089278.
PR 02-JUL-1998; 98US-0091647.
PR 24-AUG-1998; 98US-0097639.
PR 22-APR-1999; 99US-0097639.
XX
PA (ALPH-) ALPHAGENE INC.
XX
PI Valenzuela D, Yuan O, Hoffman H, Hall J, Rapiejko P;
XX
DR WPI; 2000-052801/04.
XX
PT New polynucleotides encoding secreted human proteins, derived from
PT human fetal brain, adult skin, adult brain, adult heart, adult thymus
PT and adult aorta cDNA libraries. -
XX
PS Disclosure; Page 270; 282pp; English.
XX
CC This invention describes novel human secreted proteins which are encoded
CC by polynucleotides obtained from fetal brain, adult skin, adult brain,
CC adult heart, adult thymus and adult aorta cDNA libraries. The
CC polynucleotides and proteins are predicted to have biological activities
CC which would make them suitable for treating, preventing or ameliorating
CC medical conditions in humans and animals, although no supporting data
CC is given. Suggested activities include nutritional activity, cytokine
CC and cell proliferation/differentiation activity, immune stimulating
CC (e.g. as vaccines) or suppressing activity, hematopoiesis regulating
CC activity, tissue growth activity, activin/inhibin activity,
CC chemotactic/chemokinetic activity, hemostatic and thrombolytic activity,
CC receptor/ligand activity, anti-inflammatory activity, cadherin/tumor
CC invasion suppressor activity, and tumor inhibition activity. The
CC polynucleotides are also stated to be useful for gene therapy.
CC AAZ43809-Z43840 represent DNA probes used to isolate the polynucleotides
CC represented in AAZ43777-Z43808 which encode the secreted proteins
CC represented in AAY50905-Y50947.
XX
SQ Sequence 20 BP; 4 A; 7 C; 4 G; 5 T; 0 other;

Query Match 71.0%; Score 14.2; DB 21; Length 20;
Best Local Similarity 84.2%; Pred. No. 1.8e+03;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 TGGGAATAGATGAAGTTGC 19
|||| | |||||||||| |
Db 19 TGGGCAGAGATGAAGTTCC 1

RESULT 9
ABN39186/c
ID ABN39186 standard; DNA; 60 BP.
XX
AC ABN39186;
XX
DT 15-JUL-2002 (first entry)
XX
DE Human spliced transcript detection oligonucleotide SEQ ID NO:11934.
XX
KW Human; mouse; rat; splice transcript; detection; RNA transcript;
KW splice variant; transcriptome; oligonucleotide library; ss.
XX
OS Homo sapiens.
XX
PN WO200210449-A2.
XX
PD 07-FEB-2002.
XX
PF 20-JUL-2001; 2001WO-IB01903.
XX
PR 28-JUL-2000; 2000US-221607P.
PR 02-MAY-2001; 2001US-287724P.
XX
PA (COMP-) COMPUGEN INC.
XX
PI Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
XX
DR WPI; 2002-257383/30.
XX
PT New oligonucleotide libraries comprising oligonucleotides which
PT selectively hybridize to mRNAs transcribed from a transcription unit of
PT a genome, useful for detecting tissue-, pathology-, and
PT developmental-specific genes -
XX
PS Example 1; SEQ ID 11934; 47pp; English.
XX
CC The present invention describes oligonucleotide libraries for detecting
CC messenger RNAs that populate a (sub-)transcriptome, where the
CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC transcription units that populate a genome. The library comprises
CC several oligonucleotides, each capable of hybridising selectively to a
CC set of messenger RNAs transcribed from a given transcription unit of
CC the genome, which encodes one or more messenger RNA splice variants.
CC The oligonucleotide libraries are useful for detecting mRNAs from a
CC biological sample, in expression profiling studies, in qualitatively or
CC quantitatively characterising the corresponding transcriptome, and in
CC detecting RNA transcripts and splice variants of human or animal
CC transcriptomes. The libraries may also be used as specialised mini
CC libraries to detect transcripts of a sub-transcriptome under a
CC particular biological or pathological state, and so allowing the
CC detection of tissue- and pathology-specific genes such as those genes
CC only expressed in specific tissue under a specific pathological
CC condition; to detect developmental specific genes; and to detect RNA
CC transcripts and splice variants of a transcriptome of a patient suffering
CC from a particular disorder. ABN27253 to ABN59589 represent
CC oligonucleotide sequences from rats, humans and mice, which are used in
CC the exemplification of the present invention.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 60 BP; 16 A; 12 C; 14 G; 18 T; 0 other;

Query Match 71.0%; Score 14.2; DB 24; Length 60;
Best Local Similarity 84.2%; Pred. No. 2e+03;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 TGGGAATAGATGAAGTTGC 19
||| ||||| ||||||| |
Db 37 TGGAAATAGTTGAAGTTCC 19

RESULT 10
ABN80359
ID ABN80359 standard; DNA; 26 BP.
XX
AC ABN80359;
XX
DT 16-JUL-2002 (first entry)
XX
DE PCR primer 915-F2.
XX
KW Amplification; chimeric oligonucleotide primer; disease diagnosis;
KW polymerase chain reaction; PCR; genetic engineering; blood; urine;
KW plant tissue; animal tissue; assay; soil; food; microorganism; ss.
XX
OS Thermotoga maritima.
XX
PN WO200216639-A1.
XX
PD 28-FEB-2002.
XX
PF 21-AUG-2001; 2001WO-JP07139.

XX
PR 23-AUG-2000; 2000JP-0251981.
PR 19-SEP-2000; 2000JP-0284419.
PR 22-SEP-2000; 2000JP-0288750.
PR 03-APR-2001; 2001JP-0104191.
XX
PA (TAKI) TAKARA SHUZO CO LTD.
XX
PI Sagawa H, Uemori T, Mukai H, Yamamoto J, Tomono J, Kobayashi E;
PI Enoki T, Asada K, Kato I;
XX
DR WPI; 2002-351653/38.
XX
PT Amplifying a target nucleic acid in sample, useful in e.g. clinical
PT applications, genetic engineering and for assaying blood, urine, plant
PT and animal tissues and environmental materials like soil and food -
XX
PS Examples; Page 249; 332pp; Japanese.
XX
CC The invention relates to the amplification of a nucleic acid. This
CC comprises using a nucleic acid as template, deoxypolynucleotide
CC 3-phosphate, a chimeric oligonucleotide primer with a ribonucleotide
CC provided at the 3'-terminus or in the 3'-terminal side, DNA polymerase
CC with a chain-transfer activity, an RNaseH or endonuclease, and incubating
CC the mixture to give a reaction product. The method is useful for
CC amplifying a target nucleic acid in a sample, which is useful in e.g.
CC clinical applications including disease diagnosis, genetic engineering,
CC in assaying blood, urine, plant and animal tissues, environmental
CC materials like soil and food and identification of microorganisms. The
CC method of the invention, known as an isothermal and chimeric
CC primer-initiated amplification of nucleic acids (ICAN) method, is highly
CC sensitive and specific. Sequences given in records ABN80338-ABN80532
CC represent target nucleic acids, chimeric oligonucleotides and
CC oligonucleotide primers and probes of the invention. Chimeric
CC oligonucleotides are DNA/RNA hybrids.
XX
SQ Sequence 26 BP; 10 A; 4 C; 6 G; 6 T; 0 other;

Query Match 69.0%; Score 13.8; DB 24; Length 26;
Best Local Similarity 88.2%; Pred. No. 2.9e+03;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 TGGGAATAGATGAAGTT 17
|||||||||||||| ||
Db 7 TGGGAATAGATGAGCTT 23

RESULT 11
ABL51232
ID ABL51232 standard; DNA; 26 BP.
XX
AC ABL51232;
XX
DT 27-JUN-2002 (first entry)
XX
DE Thermotolerant ribonuclease H (RNase HII) PCR primer SEQ ID NO:25.
XX
KW Thermotolerant ribonuclease H; RNase H; RNase HII; RNase HIII; enzyme;
KW genetic engineering; PCR primer; ss.
XX
OS Thermotoga maritima.
XX
PN WO200222831-A1.
XX
PD 21-MAR-2002.
XX
PF 13-SEP-2001; 2001WO-JP07930.
XX
PR 14-SEP-2000; 2000JP-0280785.
PR 07-MAR-2001; 2001JP-0064074.
XX
PA (TAKI) TAKARA SHUZO CO LTD.
XX
PI Uemori T, Sato Y, Koyama N, Hirano R, Takakura H, Kobori H;
PI Hashimoto Y, Asada K, Kato I;
XX
DR WPI; 2002-362349/39.
XX
PT Polypeptides with thermotolerant ribonuclease H activity and genes
PT encoding them for genetic engineering application -
XX
PS Example 5; Page 80; 113pp; Japanese.
XX
CC The present invention describes proteins having thermotolerant
CC ribonuclease H (RNase H) activity. The RNase H proteins and
CC polynucleotide sequences encoding them can be used for the preparation
CC of highly active thermotolerant RNase H on an industrial scale for
CC genetic engineering applications. The present sequence represents a
CC PCR primer for Thermotoga maritima RNase HII, which is used in an
CC example from the present invention.
XX
SQ Sequence 26 BP; 10 A; 4 C; 6 G; 6 T; 0 other;

Query Match 69.0%; Score 13.8; DB 24; Length 26;
Best Local Similarity 88.2%; Pred. No. 2.9e+03;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 TGGGAATAGATGAAGTT 17
|||||||||||||| ||
Db 7 TGGGAATAGATGAGCTT 23

RESULT 12
ABN80358
ID ABN80358 standard; DNA; 28 BP.
XX
AC ABN80358;
XX
DT 16-JUL-2002 (first entry)
XX
DE PCR primer 915-F1.
XX
KW Amplification; chimeric oligonucleotide primer; disease diagnosis;
KW polymerase chain reaction; PCR; genetic engineering; blood; urine;
KW plant tissue; animal tissue; assay; soil; food; microorganism; ss.
XX
OS Thermotoga maritima.
XX
PN WO200216639-A1.
XX
PD 28-FEB-2002.
XX
PF 21-AUG-2001; 2001WO-JP07139.
XX
PR 23-AUG-2000; 2000JP-0251981.
PR 19-SEP-2000; 2000JP-0284419.
PR 22-SEP-2000; 2000JP-0288750.
PR 03-APR-2001; 2001JP-0104191.
XX
PA (TAKI) TAKARA SHUZO CO LTD.
XX
PI Sagawa H, Uemori T, Mukai H, Yamamoto J, Tomono J, Kobayashi E;
PI Enoki T, Asada K, Kato I;
XX
DR WPI; 2002-351653/38.
XX
PT Amplifying a target nucleic acid in sample, useful in e.g. clinical
PT applications, genetic engineering and for assaying blood, urine, plant
PT and animal tissues and environmental materials like soil and food -
XX
PS Examples; Page 249; 332pp; Japanese.
XX
CC The invention relates to the amplification of a nucleic acid. This
CC comprises using a nucleic acid as template, deoxypolynucleotide
CC 3-phosphate, a chimeric oligonucleotide primer with a ribonucleotide
CC provided at the 3'-terminus or in the 3'-terminal side, DNA polymerase

```
CC   with a chain-transfer activity, an RNaseH or endonuclease, and incubating
CC   the mixture to give a reaction product. The method is useful for
CC   amplifying a target nucleic acid in a sample, which is useful in e.g.
CC   clinical applications including disease diagnosis, genetic engineering,
CC   in assaying blood, urine, plant and animal tissues, environmental
CC   materials like soil and food and identification of microorganisms. The
CC   method of the invention, known as an isothermal and chimeric
CC   primer-initiated amplification of nucleic acids (ICAN) method, is highly
CC   sensitive and specific. Sequences given in records ABN80338-ABN80532
CC   represent target nucleic acids, chimeric oligonucleotides and
CC   oligonucleotide primers and probes of the invention. Chimeric
CC   oligonucleotides are DNA/RNA hybrids.
XX
SQ   Sequence 28 BP; 11 A; 3 C; 7 G; 7 T; 0 other;

  Query Match           69.0%;  Score 13.8;  DB 24;  Length 28;
  Best Local Similarity  88.2%;  Pred. No. 2.9e+03;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 TGGGAATAGATGAAGTT 17
          |||||||||||||  ||
Db      9 TGGGAATAGATGAGCTT 25


RESULT 13
ABL51231
ID   ABL51231 standard; DNA; 28 BP.
XX
AC   ABL51231;
XX
DT   27-JUN-2002  (first entry)
XX
DE   Thermotolerant ribonuclease H (RNase HII) PCR primer SEQ ID NO:24.
XX
KW   Thermotolerant ribonuclease H; RNase H; RNase HII; RNase HIII; enzyme;
KW   genetic engineering; PCR primer; ss.
XX
OS   Thermotoga maritima.
XX
PN   WO200222831-A1.
XX
PD   21-MAR-2002.
XX
PF   13-SEP-2001; 2001WO-JP07930.
XX
PR   14-SEP-2000; 2000JP-0280785.
PR   07-MAR-2001; 2001JP-0064074.
XX
PA   (TAKI ) TAKARA SHUZO CO LTD.
XX
PI   Uemori T,  Sato Y,  Koyama N,  Hirano R,  Takakura H,  Kobori H;
PI   Hashimoto Y,  Asada K,  Kato I;
XX
DR   WPI; 2002-362349/39.
XX
PT   Polypeptides with thermotolerant ribonuclease H activity and genes
PT   encoding them for genetic engineering application  -
XX
PS   Example 5; Page 79; 113pp; Japanese.
XX
CC   The present invention describes proteins having thermotolerant
CC   ribonuclease H (RNase H) activity. The RNase H proteins and
CC   polynucleotide sequences encoding them can be used for the preparation
CC   of highly active thermotolerant RNase H on an industrial scale for
CC   genetic engineering applications. The present sequence represents a
CC   PCR primer for Thermotoga maritima RNase HII, which is used in an
CC   example from the present invention.
XX
SQ   Sequence 28 BP; 11 A; 3 C; 7 G; 7 T; 0 other;

  Query Match           69.0%;  Score 13.8;  DB 24;  Length 28;
  Best Local Similarity  88.2%;  Pred. No. 2.9e+03;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 TGGGAATAGATGAAGTT 17
          |||||||||||||  ||
Db      9 TGGGAATAGATGAGCTT 25


RESULT 14
AAA62774
ID   AAA62774 standard; DNA; 37 BP.
XX
AC   AAA62774;
XX
DT   25-SEP-2000  (first entry)
XX
DE   Endoglucanase PCR primer pRIN-Bgl.
XX
KW   Endoglucanase; cellulose breakdown; produce pulp; papermaking;
KW   animal foodstuff; primer; ss.
XX
OS   Synthetic.
XX
PN   WO200024879-A1.
XX
PD   04-MAY-2000.
XX
PF   25-OCT-1999;   99WO-JP05884.
XX
PR   23-OCT-1998;   98JP-0302387.
XX
PA   (MEIJ ) MEIJI SEIKA KAISHA LTD.
XX
PI   Nakamura Y,  Moriya T,  Baba Y,  Yanai K,  Sumida N,  Nishimura T;
PI   Murashima K,  Nakane A,  Yaguchi T,  Koga J,  Murakami T,  Kono T;
XX
DR   WPI; 2000-365117/31.
XX
PT   Endoglucanases of fungal origin with high activity under alkaline
PT   conditions for production of paper pulp and animal feedstuffs  -
XX
PS   Claim 51; Page 58; 180pp; Japanese.
XX
CC   This sequence represents a PCR primer used in the identification of an
CC   endoglucanase encoding protein. The invention relates to an
CC   endoglucanase of fungal origin which can completely break down purified
CC   cellulose at a concentration of less than 1mg protein/litre, and produces
CC   more than 50% breakdown of cellulose at pH 8.5. The invention includes
CC   endoglucanase protein sequences (see AAB09825-B09830), endoglucanase
CC   nucleotide sequences (see AAA62726-A62732) and primers (AAA62733-A62802)
CC   which are used in the identification of the endoglucanase sequences, and
CC   in the construction of vectors containing the polynucleotides. The
CC   endoglucanase enzymes are used for the production of pulp for papermaking
CC   and for the production of animal foodstuffs.
XX
SQ   Sequence 37 BP; 11 A; 3 C; 11 G; 12 T; 0 other;

  Query Match           69.0%;  Score 13.8;  DB 21;  Length 37;
  Best Local Similarity  88.2%;  Pred. No. 3e+03;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 TGGGAATAGATGAAGTT 17
          |||||  ||||||||||
Db     10 TGGGACAAGATGAAGTT 26


RESULT 15
ABN48099
ID   ABN48099 standard; DNA; 60 BP.
XX
AC   ABN48099;
XX
DT   15-JUL-2002  (first entry)
XX
DE   Human spliced transcript detection oligonucleotide SEQ ID NO:20847.
```

XX
KW Human; mouse; rat; splice transcript; detection; RNA transcript;
KW splice variant; transcriptome; oligonucleotide library; ss.
XX
OS Homo sapiens.
XX
PN WO200210449-A2.
XX
PD 07-FEB-2002.
XX
PF 20-JUL-2001; 2001WO-IB01903.
XX
PR 28-JUL-2000; 2000US-221607P.
PR 02-MAY-2001; 2001US-287724P.
XX
PA (COMP-) COMPUGEN INC.
XX
PI Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
XX
DR WPI; 2002-257383/30.
XX
PT New oligonucleotide libraries comprising oligonucleotides which
PT selectively hybridize to mRNAs transcribed from a transcription unit of
PT a genome, useful for detecting tissue-, pathology-, and
PT developmental-specific genes -
XX
PS Example 1; SEQ ID 20847; 47pp; English.
XX
CC The present invention describes oligonucleotide libraries for detecting
CC messenger RNAs that populate a (sub-)transcriptome, where the
CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC transcription units that populate a genome. The library comprises
CC several oligonucleotides, each capable of hybridising selectively to a
CC set of messenger RNAs transcribed from a given transcription unit of
CC the genome, which encodes one or more messenger RNA splice variants.
CC The oligonucleotide libraries are useful for detecting mRNAs from a
CC biological sample, in expression profiling studies, in qualitatively or
CC quantitatively characterising the corresponding transcriptome, and in
CC detecting RNA transcripts and splice variants of human or animal
CC transcriptomes. The libraries may also be used as specialised mini
CC libraries to detect transcripts of a sub-transcriptome under a
CC particular biological or pathological state, and so allowing the
CC detection of tissue- and pathology-specific genes such as those genes
CC only expressed in specific tissue under a specific pathological
CC condition; to detect developmental specific genes; and to detect RNA
CC transcripts and splice variants of a transcriptome of a patient suffering
CC from a particular disorder. ABN27253 to ABN59589 represent
CC oligonucleotide sequences from rats, humans and mice, which are used in
CC the exemplification of the present invention.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 60 BP; 19 A; 12 C; 16 G; 13 T; 0 other;

Query Match 69.0%; Score 13.8; DB 24; Length 60;
Best Local Similarity 88.2%; Pred. No. 3.2e+03;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy      2 GGGAATAGATGAAGTTG 18
          ||||||||  |||||  ||
Db      2 GGGAATAGGTGAAGGTG 18
```

Search completed: November 23, 2002, 06:29:15
Job time : 101.6 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 05:54:41 ; Search time 17.25 Seconds
(without alignments)
439.108 Million cell updates/sec

Title: US-09-296-264-15
Perfect score: 20
Sequence: 1 tgggaatagatgaagttgcc 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 335578 seqs, 189365133 residues

Total number of hits satisfying chosen parameters: 195614

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	14.4	72.0	60	10	US-09-934-465-5	Sequence 5, Appli
c	2	14.4	72.0	60	10	US-09-884-733-8	Sequence 8, Appli
	3	13.8	69.0	89	10	US-09-764-877-357	Sequence 357, App
c	4	13.8	69.0	100	10	US-09-728-445-277	Sequence 277, App
c	5	13.4	67.0	78	10	US-09-864-761-23951	Sequence 23951, A
	6	13	65.0	17	10	US-09-866-108-6732	Sequence 6732, Ap
	7	13	65.0	17	10	US-09-866-108-6733	Sequence 6733, Ap
	8	13	65.0	17	10	US-09-866-108-6734	Sequence 6734, Ap
	9	13	65.0	17	10	US-09-866-108-6735	Sequence 6735, Ap
	10	13	65.0	17	10	US-09-866-108-6736	Sequence 6736, Ap
	11	13	65.0	25	10	US-09-866-108-11624	Sequence 11624, A
	12	13	65.0	25	10	US-09-866-108-11625	Sequence 11625, A
	13	13	65.0	25	10	US-09-866-108-11626	Sequence 11626, A
	14	13	65.0	25	10	US-09-866-108-11627	Sequence 11627, A
	15	13	65.0	25	10	US-09-866-108-11628	Sequence 11628, A
	16	13	65.0	25	10	US-09-866-108-11629	Sequence 11629, A
	17	13	65.0	25	10	US-09-866-108-11630	Sequence 11630, A
	18	13	65.0	25	10	US-09-866-108-11631	Sequence 11631, A
	19	13	65.0	25	10	US-09-866-108-11632	Sequence 11632, A
	20	13	65.0	25	10	US-09-866-108-11633	Sequence 11633, A
	21	13	65.0	25	10	US-09-866-108-11634	Sequence 11634, A
	22	13	65.0	25	10	US-09-866-108-11635	Sequence 11635, A
	23	13	65.0	25	10	US-09-866-108-11636	Sequence 11636, A
c	24	12.8	64.0	31	10	US-09-801-274-670	Sequence 670, App
c	25	12.8	64.0	93	10	US-09-864-761-26178	Sequence 26178, A
	26	12.6	63.0	77	10	US-09-864-761-33023	Sequence 33023, A
c	27	12.6	63.0	93	10	US-09-294-093B-902	Sequence 902, App
c	28	12.2	61.0	40	10	US-09-757-207-11	Sequence 11, Appl
c	29	12.2	61.0	86	10	US-09-864-761-27450	Sequence 27450, A
c	30	12.2	61.0	94	10	US-09-864-761-33219	Sequence 33219, A
	31	12	60.0	17	10	US-09-866-108-6731	Sequence 6731, Ap
	32	12	60.0	17	10	US-09-866-108-6737	Sequence 6737, Ap
	33	12	60.0	25	10	US-09-866-108-11623	Sequence 11623, A
	34	12	60.0	25	10	US-09-866-108-11637	Sequence 11637, A
c	35	12	60.0	31	10	US-09-801-274-1096	Sequence 1096, Ap
	36	12	60.0	50	9	US-09-978-295A-356	Sequence 356, App
	37	12	60.0	50	9	US-09-992-598-187	Sequence 187, App
	38	12	60.0	50	9	US-09-978-697-356	Sequence 356, App
	39	12	60.0	50	10	US-09-989-722-187	Sequence 187, App
	40	12	60.0	50	10	US-09-989-723-187	Sequence 187, App
	41	12	60.0	50	10	US-09-989-279-187	Sequence 187, App
	42	12	60.0	50	10	US-09-989-727-187	Sequence 187, App
	43	12	60.0	50	10	US-09-989-731-187	Sequence 187, App
	44	12	60.0	50	10	US-09-989-732-187	Sequence 187, App
	45	12	60.0	50	10	US-09-991-073-187	Sequence 187, App

ALIGNMENTS

RESULT 1
US-09-934-465-5/c
; Sequence 5, Application US/09934465
; Patent No. US20020102233A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; TITLE OF INVENTION: APO-2 LIGAND
; FILE REFERENCE: 11669.22US03
; CURRENT APPLICATION NUMBER: US/09/934,465
; CURRENT FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: 08/584,031
; PRIOR FILING DATE: 1996-01-09
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 60
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(60)
; OTHER INFORMATION: Sequence is synthesized
US-09-934-465-5

Query Match 72.0%; Score 14.4; DB 10; Length 60;
Best Local Similarity 93.8%; Pred. No. 2e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy      1 TGGGAATAGATGAAGT 16
          |||||||||||| |||
Db     47 TGGGAATAGATGTAGT 32
```

RESULT 2
US-09-884-733-8/c
; Sequence 8, Application US/09884733
; Patent No. US20020123116A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; TITLE OF INVENTION: Apo-2 Ligand Inhibitor
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:

```
;             ADDRESSEE: Genentech, Inc.
;             STREET: 460 Point San Bruno Blvd
;             CITY: South San Francisco
;             STATE: California
;             COUNTRY: USA
;             ZIP: 94080
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: WinPatin (Genentech)
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/884,733
;             FILING DATE: 19-Jun-2001
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 09/304,003
;             FILING DATE: 14-JUNE-2000
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Marschang, Diane L.
;             REGISTRATION NUMBER: 35,600
;             REFERENCE/DOCKET NUMBER: P1007
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 415/225-5416
;             TELEFAX: 415/952-9881
;             TELEX: 910/371-7168
;   INFORMATION FOR SEQ ID NO: 8:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 60 base pairs
;             TYPE: Nucleic Acid
;             STRANDEDNESS: Single
;             TOPOLOGY: Linear
;        SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-09-884-733-8

  Query Match             72.0%;  Score 14.4;  DB 10;  Length 60;
  Best Local Similarity   93.8%;  Pred. No. 2e+02;
  Matches   15;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 TGGGAATAGATGAAGT 16
           ||||||||||| |||
Db      47 TGGGAATAGATGTAGT 32


RESULT 3
US-09-764-877-357
; Sequence 357, Application US/09764877
; Patent No. US20020147140A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;  FILE REFERENCE: PC005
;  CURRENT APPLICATION NUMBER: US/09/764,877
;  CURRENT FILING DATE:  2001-01-17
;  Prior application data removed - refer to PALM or file wrapper
;  NUMBER OF SEQ ID NOS: 4031
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 357
;   LENGTH: 89
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-764-877-357

  Query Match             69.0%;  Score 13.8;  DB 10;  Length 89;
  Best Local Similarity   88.2%;  Pred. No. 4.3e+02;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 TGGGAATAGATGAAGTT 17
           |||||| ||||||||  |
Db      29 TGGGAAAAGATGAAGAT 45


RESULT 4
US-09-728-445-277/c
; Sequence 277, Application US/09728445
; Patent No. US20020102543A1
; GENERAL INFORMATION:
;  APPLICANT: Friedrich, Glenn
;  APPLICANT:  Zambrowicz, Brian
;  APPLICANT:  Sands, Arthur T.
;  TITLE OF INVENTION: No. US20020102543A1el Mutated Mammalian Cells and
;  TITLE OF INVENTION:  Animals
;  FILE REFERENCE: LEX-0102-USA
;  CURRENT APPLICATION NUMBER: US/09/728,445
;  CURRENT FILING DATE:  2000-11-30
;  PRIOR APPLICATION NUMBER: US 60/168,358
;  PRIOR FILING DATE: 1999-12-01
;  NUMBER OF SEQ ID NOS: 891
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 277
;   LENGTH: 100
;   TYPE: DNA
;   ORGANISM: Mus musculus
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(100)
;   OTHER INFORMATION: n = A,T,C or G
US-09-728-445-277

  Query Match             69.0%;  Score 13.8;  DB 10;  Length 100;
  Best Local Similarity   83.3%;  Pred. No. 4.4e+02;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 TGGGAATAGATGAAGTTG 18
           ||||| ||||  ||||||
Db      62 TGGGACTAGAANAAGTTG 45


RESULT 5
US-09-864-761-23951/c
; Sequence 23951, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
```

```
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 23951
;   LENGTH: 78
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AL117334:29
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.6
;   OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.8
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.7
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3.3
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.9
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.7
;   OTHER INFORMATION: NT HIT: M20231.1, EVALUE 2.40e+00
;   OTHER INFORMATION: EST_HUMAN HIT: AW103658.1, EVALUE 1.00e-36
US-09-864-761-23951

  Query Match            67.0%;  Score 13.4;  DB 10;  Length 78;
  Best Local Similarity  93.3%;  Pred. No. 6.7e+02;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 GGGAATAGATGAAGT 16
           ||||||| ||||||
Db      15 GGGAATAGCTGAAGT 1


RESULT 6
US-09-866-108-6732
; Sequence 6732, Application US/09866108
; Patent No. US20020048800A1
; GENERAL INFORMATION:
;  APPLICANT: GU, Yizhong
;  APPLICANT:  JI, Yonggang
;  APPLICANT:  PENN, Sharron G.
;  APPLICANT:  HANZEL, David K.
;  APPLICANT:  RANK, David R.
;  APPLICANT:  CHEN, Wensheng
;  APPLICANT:  SHANNON, Mark
;  TITLE OF INVENTION: MYOSIN-LIKE GENE EXPRESSED IN HUMAN HEART AND MUSCLE
;  FILE REFERENCE: AEOMICA-7
;  CURRENT APPLICATION NUMBER: US/09/866,108
;  CURRENT FILING DATE:  2001-05-25
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 60/266,860
;  PRIOR FILING DATE: 2001-02-05
;  NUMBER OF SEQ ID NOS: 15752
;  SOFTWARE: Aeomica Sequence Listing Engine
; SEQ ID NO 6732
;   LENGTH: 17
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-866-108-6732

  Query Match            65.0%;  Score 13;  DB 10;  Length 17;
  Best Local Similarity  100.0%;  Pred. No. 7.9e+02;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 GGGAATAGATGAA 14
           |||||||||||||
Db       5 GGGAATAGATGAA 17


RESULT 7
US-09-866-108-6733
; Sequence 6733, Application US/09866108
; Patent No. US20020048800A1
; GENERAL INFORMATION:
;  APPLICANT: GU, Yizhong
;  APPLICANT:  JI, Yonggang
;  APPLICANT:  PENN, Sharron G.
;  APPLICANT:  HANZEL, David K.
;  APPLICANT:  RANK, David R.
;  APPLICANT:  CHEN, Wensheng
;  APPLICANT:  SHANNON, Mark
;  TITLE OF INVENTION: MYOSIN-LIKE GENE EXPRESSED IN HUMAN HEART AND MUSCLE
;  FILE REFERENCE: AEOMICA-7
;  CURRENT APPLICATION NUMBER: US/09/866,108
;  CURRENT FILING DATE:  2001-05-25
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
```

```
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 60/266,860
;  PRIOR FILING DATE: 2001-02-05
;  NUMBER OF SEQ ID NOS: 15752
;  SOFTWARE: Aeomica Sequence Listing Engine
; SEQ ID NO 6733
;   LENGTH: 17
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-866-108-6733

  Query Match            65.0%;  Score 13;  DB 10;  Length 17;
  Best Local Similarity  100.0%;  Pred. No. 7.9e+02;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      2 GGGAATAGATGAA 14
          |||||||||||||
Db      4 GGGAATAGATGAA 16


RESULT 8
US-09-866-108-6734
; Sequence 6734, Application US/09866108
; Patent No. US20020048800A1
; GENERAL INFORMATION:
;  APPLICANT: GU, Yizhong
;  APPLICANT:  JI, Yonggang
;  APPLICANT:  PENN, Sharron G.
;  APPLICANT:  HANZEL, David K.
;  APPLICANT:  RANK, David R.
;  APPLICANT:  CHEN, Wensheng
;  APPLICANT:  SHANNON, Mark
;  TITLE OF INVENTION: MYOSIN-LIKE GENE EXPRESSED IN HUMAN HEART AND MUSCLE
;  FILE REFERENCE: AEOMICA-7
;  CURRENT APPLICATION NUMBER: US/09/866,108
;  CURRENT FILING DATE:  2001-05-25
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 60/266,860
;  PRIOR FILING DATE: 2001-02-05
;  NUMBER OF SEQ ID NOS: 15752
;  SOFTWARE: Aeomica Sequence Listing Engine
; SEQ ID NO 6734
;   LENGTH: 17
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-866-108-6734

  Query Match            65.0%;  Score 13;  DB 10;  Length 17;
  Best Local Similarity  100.0%;  Pred. No. 7.9e+02;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      2 GGGAATAGATGAA 14
          |||||||||||||
Db      3 GGGAATAGATGAA 15


RESULT 9
US-09-866-108-6735
; Sequence 6735, Application US/09866108
; Patent No. US20020048800A1
; GENERAL INFORMATION:
;  APPLICANT: GU, Yizhong
;  APPLICANT:  JI, Yonggang
;  APPLICANT:  PENN, Sharron G.
;  APPLICANT:  HANZEL, David K.
;  APPLICANT:  RANK, David R.
;  APPLICANT:  CHEN, Wensheng
;  APPLICANT:  SHANNON, Mark
;  TITLE OF INVENTION: MYOSIN-LIKE GENE EXPRESSED IN HUMAN HEART AND MUSCLE
;  FILE REFERENCE: AEOMICA-7
;  CURRENT APPLICATION NUMBER: US/09/866,108
;  CURRENT FILING DATE:  2001-05-25
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 60/266,860
;  PRIOR FILING DATE: 2001-02-05
;  NUMBER OF SEQ ID NOS: 15752
;  SOFTWARE: Aeomica Sequence Listing Engine
; SEQ ID NO 6735
;   LENGTH: 17
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-866-108-6735

  Query Match            65.0%;  Score 13;  DB 10;  Length 17;
  Best Local Similarity  100.0%;  Pred. No. 7.9e+02;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      2 GGGAATAGATGAA 14
          |||||||||||||
Db      2 GGGAATAGATGAA 14
```

```
RESULT 10
US-09-866-108-6736
; Sequence 6736, Application US/09866108
; Patent No. US20020048800A1
; GENERAL INFORMATION:
;  APPLICANT: GU, Yizhong
;  APPLICANT:  JI, Yonggang
;  APPLICANT:  PENN, Sharron G.
;  APPLICANT:  HANZEL, David K.
;  APPLICANT:  RANK, David R.
;  APPLICANT:  CHEN, Wensheng
;  APPLICANT:  SHANNON, Mark
;  TITLE OF INVENTION: MYOSIN-LIKE GENE EXPRESSED IN HUMAN HEART AND MUSCLE
;  FILE REFERENCE: AEOMICA-7
;  CURRENT APPLICATION NUMBER: US/09/866,108
;  CURRENT FILING DATE:  2001-05-25
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 60/266,860
;  PRIOR FILING DATE: 2001-02-05
;  NUMBER OF SEQ ID NOS: 15752
;  SOFTWARE: Aeomica Sequence Listing Engine
; SEQ ID NO 6736
;   LENGTH: 17
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-866-108-6736

  Query Match             65.0%;  Score 13;  DB 10;  Length 17;
  Best Local Similarity   100.0%;  Pred. No. 7.9e+02;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 GGGAATAGATGAA 14
           |||||||||||||
Db       1 GGGAATAGATGAA 13


RESULT 11
US-09-866-108-11624
; Sequence 11624, Application US/09866108
; Patent No. US20020048800A1
; GENERAL INFORMATION:
;  APPLICANT:  GU, Yizhong
;  APPLICANT:  JI, Yonggang
;  APPLICANT:  PENN, Sharron G.
;  APPLICANT:  HANZEL, David K.
;  APPLICANT:  RANK, David R.
;  APPLICANT:  CHEN, Wensheng
;  APPLICANT:  SHANNON, Mark
;  TITLE OF INVENTION: MYOSIN-LIKE GENE EXPRESSED IN HUMAN HEART AND MUSCLE
;  FILE REFERENCE: AEOMICA-7
;  CURRENT APPLICATION NUMBER: US/09/866,108
;  CURRENT FILING DATE:  2001-05-25
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 60/266,860
;  PRIOR FILING DATE: 2001-02-05
;  NUMBER OF SEQ ID NOS: 15752
;  SOFTWARE: Aeomica Sequence Listing Engine
; SEQ ID NO 11624
;   LENGTH: 25
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-866-108-11624

  Query Match             65.0%;  Score 13;  DB 10;  Length 25;
  Best Local Similarity   100.0%;  Pred. No. 8.5e+02;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 GGGAATAGATGAA 14
           |||||||||||||
Db      13 GGGAATAGATGAA 25


RESULT 12
US-09-866-108-11625
; Sequence 11625, Application US/09866108
; Patent No. US20020048800A1
; GENERAL INFORMATION:
;  APPLICANT: GU, Yizhong
;  APPLICANT:  JI, Yonggang
;  APPLICANT:  PENN, Sharron G.
;  APPLICANT:  HANZEL, David K.
;  APPLICANT:  RANK, David R.
;  APPLICANT:  CHEN, Wensheng
;  APPLICANT:  SHANNON, Mark
;  TITLE OF INVENTION: MYOSIN-LIKE GENE EXPRESSED IN HUMAN HEART AND MUSCLE
;  FILE REFERENCE: AEOMICA-7
;  CURRENT APPLICATION NUMBER: US/09/866,108
;  CURRENT FILING DATE:  2001-05-25
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
```

```
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 60/266,860
;  PRIOR FILING DATE: 2001-02-05
;  NUMBER OF SEQ ID NOS: 15752
;  SOFTWARE: Aeomica Sequence Listing Engine
; SEQ ID NO 11625
;   LENGTH: 25
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-866-108-11625

  Query Match             65.0%;  Score 13;  DB 10;  Length 25;
  Best Local Similarity   100.0%;  Pred. No. 8.5e+02;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 GGGAATAGATGAA 14
           |||||||||||||
Db      12 GGGAATAGATGAA 24


RESULT 13
US-09-866-108-11626
; Sequence 11626, Application US/09866108
; Patent No. US20020048800A1
; GENERAL INFORMATION:
;  APPLICANT: GU, Yizhong
;  APPLICANT:  JI, Yonggang
;  APPLICANT:  PENN, Sharron G.
;  APPLICANT:  HANZEL, David K.
;  APPLICANT:  RANK, David R.
;  APPLICANT:  CHEN, Wensheng
;  APPLICANT:  SHANNON, Mark
;  TITLE OF INVENTION: MYOSIN-LIKE GENE EXPRESSED IN HUMAN HEART AND MUSCLE
;  FILE REFERENCE: AEOMICA-7
;  CURRENT APPLICATION NUMBER: US/09/866,108
;  CURRENT FILING DATE:  2001-05-25
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 60/266,860
;  PRIOR FILING DATE: 2001-02-05
;  NUMBER OF SEQ ID NOS: 15752
;  SOFTWARE: Aeomica Sequence Listing Engine
; SEQ ID NO 11626
;   LENGTH: 25
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-866-108-11626

  Query Match             65.0%;  Score 13;  DB 10;  Length 25;
  Best Local Similarity   100.0%;  Pred. No. 8.5e+02;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 GGGAATAGATGAA 14
           |||||||||||||
Db      11 GGGAATAGATGAA 23


RESULT 14
US-09-866-108-11627
; Sequence 11627, Application US/09866108
; Patent No. US20020048800A1
; GENERAL INFORMATION:
;  APPLICANT: GU, Yizhong
;  APPLICANT:  JI, Yonggang
;  APPLICANT:  PENN, Sharron G.
;  APPLICANT:  HANZEL, David K.
;  APPLICANT:  RANK, David R.
;  APPLICANT:  CHEN, Wensheng
;  APPLICANT:  SHANNON, Mark
;  TITLE OF INVENTION: MYOSIN-LIKE GENE EXPRESSED IN HUMAN HEART AND MUSCLE
;  FILE REFERENCE: AEOMICA-7
;  CURRENT APPLICATION NUMBER: US/09/866,108
;  CURRENT FILING DATE:  2001-05-25
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
```

```
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 60/266,860
;  PRIOR FILING DATE: 2001-02-05
;  NUMBER OF SEQ ID NOS: 15752
;  SOFTWARE: Aeomica Sequence Listing Engine
; SEQ ID NO 11627
;    LENGTH: 25
;    TYPE: DNA
;    ORGANISM: Homo sapiens
US-09-866-108-11627

  Query Match            65.0%;  Score 13;  DB 10;  Length 25;
  Best Local Similarity  100.0%;  Pred. No. 8.5e+02;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 GGGAATAGATGAA 14
           |||||||||||||
Db      10 GGGAATAGATGAA 22


RESULT 15
US-09-866-108-11628
; Sequence 11628, Application US/09866108
; Patent No. US20020048800A1
; GENERAL INFORMATION:
;  APPLICANT: GU, Yizhong
;  APPLICANT:  JI, Yonggang
;  APPLICANT:  PENN, Sharron G.
;  APPLICANT:  HANZEL, David K.
;  APPLICANT:  RANK, David R.
;  APPLICANT:  CHEN, Wensheng
;  APPLICANT:  SHANNON, Mark
;  TITLE OF INVENTION: MYOSIN-LIKE GENE EXPRESSED IN HUMAN HEART AND MUSCLE
;  FILE REFERENCE: AEOMICA-7
;  CURRENT APPLICATION NUMBER: US/09/866,108
;  CURRENT FILING DATE:  2001-05-25
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 60/266,860
;  PRIOR FILING DATE: 2001-02-05
;  NUMBER OF SEQ ID NOS: 15752
;  SOFTWARE: Aeomica Sequence Listing Engine
; SEQ ID NO 11628
;    LENGTH: 25
;    TYPE: DNA
;    ORGANISM: Homo sapiens
US-09-866-108-11628

  Query Match            65.0%;  Score 13;  DB 10;  Length 25;
  Best Local Similarity  100.0%;  Pred. No. 8.5e+02;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 GGGAATAGATGAA 14
           |||||||||||||
Db       9 GGGAATAGATGAA 21


Search completed: November 23, 2002, 06:42:12
Job time : 18.25 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 25, 2002, 09:10:06 ; Search time 755.55 Seconds
(without alignments)
428.707 Million cell updates/sec

Title: US-09-296-264-15
Perfect score: 20
Sequence: 1 tgggaatagatgaagttgcc 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 357874

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	15.2	76.0	98	17	AZ922098	AZ922098 HRCot3A09
c	2	15.2	76.0	100	17	AZ922096	AZ922096 HRCot3F01
	3	14.8	74.0	64	12	BF219620	BF219620 SWOvL3CAN
c	4	14.2	71.0	46	9	AA722219	AA722219 zh21a09.s
	5	14.2	71.0	51	14	D19120	D19120 MUSGS01332
	6	14	70.0	95	17	BH813380	BH813380 SALK_0640
	7	13.8	69.0	81	10	AW432479	AW432479 sh74f05.y
	8	13.6	68.0	79	9	AA737989	AA737989 nx11f12.s
	9	13.6	68.0	90	9	AI680520	AI680520 tw82a01.x
c	10	13.6	68.0	91	12	BF465255	BF465255 UI-M-CG0p
c	11	13.6	68.0	93	14	W80226	W80226 me89f10.r1
c	12	13.6	68.0	100	12	BG063371	BG063371 H3006E03-
	13	13.4	67.0	49	9	AI453782	AI453782 tj23h02.x
c	14	13.4	67.0	67	17	TA150D05Q	AL472813 T. brucei
c	15	13.4	67.0	73	13	BI452616	BI452616 603169810
c	16	13.2	66.0	52	17	AZ919577	AZ919577 1006015G0
c	17	13.2	66.0	63	17	B33910	B33910 HS-1023-B1-
	18	13.2	66.0	73	9	AI471427	AI471427 tm11c07.x
c	19	13.2	66.0	79	9	AA218699	AA218699 zq96h07.r
c	20	13.2	66.0	82	17	AZ919576	AZ919576 1006015G0
c	21	13.2	66.0	89	17	BH789306	BH789306 SALK_0190
c	22	13.2	66.0	93	17	BH814614	BH814614 SALK_0667
	23	13.2	66.0	94	9	AA916311	AA916311 on64a10.s
c	24	13.2	66.0	96	9	AA075501	AA075501 zm87e11.r
	25	13.2	66.0	96	9	AU255664	AU255664 AU255664
c	26	13.2	66.0	96	17	AL771837	AL771837 Arabidops
	27	13.2	66.0	98	9	AI735923	AI735923 sb20f05.y
c	28	13.2	66.0	100	10	AW424747	AW424747 707064C09
c	29	13	65.0	57	9	AI930814	AI930814 sc47d10.y
	30	12.8	64.0	46	17	AZ807429	AZ807429 2M0070I18
	31	12.8	64.0	52	14	N44058	N44058 yy30c08.r1
c	32	12.8	64.0	77	13	BI451148	BI451148 kx11c12.y
c	33	12.8	64.0	80	9	AI759253	AI759253 EtESTea26
	34	12.8	64.0	88	17	BH809860	BH809860 SALK_0064
c	35	12.8	64.0	90	14	BQ754485	BQ754485 EBed01_SQ
	36	12.8	64.0	93	12	BG232720	BG232720 pa79e04.y
c	37	12.8	64.0	93	17	HSMC35E11	X88121 H.sapiens D
	38	12.8	64.0	95	9	AA822431	AA822431 vw37f09.r
c	39	12.8	64.0	95	9	AA405980	AA405980 zu67e11.s
	40	12.8	64.0	98	10	BE672882	BE672882 7a57g11.x
	41	12.6	63.0	43	9	AU256488	AU256488 AU256488
c	42	12.6	63.0	43	17	AZ616883	AZ616883 1M0446G19
c	43	12.6	63.0	50	9	AU102458	AU102458 AU102458
c	44	12.6	63.0	52	9	AA856040	AA856040 vw82a06.r
c	45	12.6	63.0	52	12	BF633988	BF633988 NF072G12D

ALIGNMENTS

RESULT 1
AZ922098/c
LOCUS AZ922098 98 bp DNA linear GSS 07-JUN-2002
DEFINITION HRCot3A09 Sorghum bicolor HRCot Sorghum bicolor genomic similar to Sorghum bicolor Retrosor-6 retroelement LTR, DNA sequence.
ACCESSION AZ922098
VERSION AZ922098.1 GI:13400381
KEYWORDS GSS.
SOURCE sorghum.
ORGANISM Sorghum bicolor
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE 1 (bases 1 to 98)
AUTHORS Peterson,D.G., Schulze,S.R., Sciara,E.B., Lee,S.A., Bowers,J.E., Nagel,A., Jiang,N., Tibbitts,D.C., Wessler,S.R. and Paterson,A.H.
TITLE Integration of Cot analysis, DNA cloning, and high-throughput sequencing facilitates genome characterization and gene discovery
JOURNAL Genome Res. 12 (5), 795-807 (2002)
MEDLINE 21992826
COMMENT Contact: Peterson DG
Plant Genome Mapping Laboratory
University of Georgia
Room 162, Riverbend Research Bldg., 110 Riverbend Rd., Athens, GA 30602, USA
Tel: 706-583-0167
Fax: 706-583-0160
Email: dgp@arches.uga.edu
Class: Hydroxyapatite-fractionated DNA.

```
FEATURES             Location/Qualifiers
     source          1. .98
                     /organism="Sorghum bicolor"
                     /cultivar="BTx623"
                     /db_xref="taxon:4558"
                     /clone_lib="Sorghum bicolor HRCot"
                     /tissue_type="leaves"
                     /dev_stage="seedling"
                     /note="Vector: pGEM-TA-Easy; A Cot analysis was performed
                     for the sorghum genome.  Based on the resulting Cot curve,
                     hydroxyapatite chromatography was used to isolate
                     'highly-repetitive' (HR), 'moderately-repetitive' (MR),
                     and 'single/low-copy' (SL) sequence components from
                     sheared genomic DNA.  The three repetition-based DNA
                     components were cloned into E. coli to produce HRCot,
                     MRCot, and SLCot genomic libraries.  Blotting and
                     sequencing data indicates that each library is
                     representative of the component from which it was derived.
                     Putative ID listings given for sequences are based on
                     comparison (blastn) with sequences in the NCBI Nr
                     Database.  Only the primary match is given (all primary E
                     values are < or = 1.00E-5).  In no instance does a 'Cot
                     clone' contain the complete sequence of its putative Nr
                     match."
BASE COUNT       29 a     23 c     14 g     32 t
ORIGIN

  Query Match             76.0%;  Score 15.2;  DB 17;  Length 98;
  Best Local Similarity   85.0%;  Pred. No. 5.8e+03;
  Matches   17;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 TGGGAATAGATGAAGTTGCC 20
          |||||| ||| |||||| ||
Db     33 TGGGAAGAGAGGAAGTTTCC 14


RESULT 2
AZ922096/c
LOCUS       AZ922096                 100 bp    DNA     linear   GSS 07-JUN-2002
DEFINITION  HRCot3F01 Sorghum bicolor HRCot Sorghum bicolor genomic similar to
            Sorghum bicolor Retrosor-6 retroelement LTR, DNA sequence.
ACCESSION   AZ922096
VERSION     AZ922096.1  GI:13400379
KEYWORDS    GSS.
SOURCE      sorghum.
  ORGANISM  Sorghum bicolor
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
            clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE   1  (bases 1 to 100)
  AUTHORS   Peterson,D.G., Schulze,S.R., Sciara,E.B., Lee,S.A., Bowers,J.E.,
            Nagel,A., Jiang,N., Tibbitts,D.C., Wessler,S.R. and Paterson,A.H.
  TITLE     Integration of Cot analysis, DNA cloning, and high-throughput
            sequencing facilitates genome characterization and gene discovery
  JOURNAL   Genome Res. 12 (5), 795-807 (2002)
  MEDLINE   21992826
COMMENT     Contact: Peterson DG
            Plant Genome Mapping Laboratory
            University of Georgia
            Room 162, Riverbend Research Bldg., 110 Riverbend Rd., Athens, GA
            30602, USA
            Tel: 706-583-0167
            Fax: 706-583-0160
            Email: dgp@arches.uga.edu
            Class: Hydroxyapatite-fractionated DNA.
FEATURES             Location/Qualifiers
     source          1. .100
                     /organism="Sorghum bicolor"
                     /cultivar="BTx623"
                     /db_xref="taxon:4558"
                     /clone_lib="Sorghum bicolor HRCot"
                     /tissue_type="leaves"
                     /dev_stage="seedling"
                     /note="Vector: pGEM-TA-Easy; A Cot analysis was performed
                     for the sorghum genome.  Based on the resulting Cot curve,
                     hydroxyapatite chromatography was used to isolate
                     'highly-repetitive' (HR), 'moderately-repetitive' (MR),
                     and 'single/low-copy' (SL) sequence components from
                     sheared genomic DNA.  The three repetition-based DNA
                     components were cloned into E. coli to produce HRCot,
                     MRCot, and SLCot genomic libraries.  Blotting and
                     sequencing data indicates that each library is
                     representative of the component from which it was derived.
                     Putative ID listings given for sequences are based on
                     comparison (blastn) with sequences in the NCBI Nr
                     Database.  Only the primary match is given (all primary E
                     values are < or = 1.00E-5).  In no instance does a 'Cot
                     clone' contain the complete sequence of its putative Nr
                     match."
BASE COUNT       37 a     20 c     11 g     32 t
ORIGIN

  Query Match             76.0%;  Score 15.2;  DB 17;  Length 100;
  Best Local Similarity   85.0%;  Pred. No. 5.8e+03;
  Matches   17;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 TGGGAATAGATGAAGTTGCC 20
          |||||| ||| |||||| ||
Db     85 TGGGAAGAGAGGAAGTTTCC 66


RESULT 3
BF219620
LOCUS       BF219620                 64 bp    mRNA    linear   EST 06-NOV-2000
DEFINITION  SWOvL3CAN72G02SK Onchocerca volvulus infective larva cDNA
            (SAW94WL-OvL3) Onchocerca volvulus cDNA clone SWOvL3CAN72G02 5',
            mRNA sequence.
ACCESSION   BF219620
VERSION     BF219620.1  GI:11118013
KEYWORDS    EST.
SOURCE      Onchocerca volvulus.
  ORGANISM  Onchocerca volvulus
            Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
            Onchocercidae; Onchocerca.
REFERENCE   1  (bases 1 to 64)
  AUTHORS   Williams,S.A., Lu,W., Lizotte-Waniewski,M. and Laney,S.J.
  TITLE     Genes expressed in infective third stage larvae of Onchocerca
            volvulus
  JOURNAL   Unpublished (1995)
COMMENT     Contact: Steven A. Williams
            Molecular Parasitology
            Smith College Department of Biological Sciences
            Department of Biological Sciences, Clark Science Center, Smith
            College, Northampton, MA, 01063, USA
            Tel: 4135853826
            Fax: 4135853786
            Email: genome@smith.edu
            Seq primer: pBluescript SK.
FEATURES             Location/Qualifiers
     source          1. .64
                     /organism="Onchocerca volvulus"
                     /strain="Sierra Leone"
                     /db_xref="taxon:6282"
                     /clone="SWOvL3CAN72G02"
                     /clone_lib="Onchocerca volvulus infective larva cDNA
                     (SAW94WL-OvL3)"
                     /lab_host="XL1-Blue MRF'"
                     /note="Vector: lambda UniZap XR; Site_1: EcoR I; Site_2:
                     Xho I; Cutaneous filarial nematode parasite of humans.
                     mRNA was prepared from third stage infective larvae of
                     Onchocerca volvulus isolated from mosquitoes 10 days after
                     infection and converted to double stranded cDNA using
                     reverse transcriptase and oligo(dT) followed by RNase H
                     and DNApol I. The library had 1.8 x 10E5 independent
                     recombinants and average insert size was 900 base pairs.
                     The library was constructed by Wenhong Lu. The library is
```

available from Dr. S.A. Williams, email genome@smith.edu."
BASE COUNT 24 a 3 c 17 g 20 t
ORIGIN

Query Match 74.0%; Score 14.8; DB 12; Length 64;
Best Local Similarity 88.9%; Pred. No. 7.9e+03;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy      1 TGGGAATAGATGAAGTTG 18
          |||||| |||||| ||||
Db     23 TGGGAACAGATGATGTTG 40
```

RESULT 4
AA722219/c
LOCUS AA722219 46 bp mRNA linear EST 02-JAN-1998
DEFINITION zh21a09.s1 Soares_pineal_gland_N3HPG Homo sapiens cDNA clone IMAGE:412696 3' similar to TR:P97356 P97356 SERINE C-PALMITOYLTRANSFERASE ;, mRNA sequence.
ACCESSION AA722219
VERSION AA722219.1 GI:2739926
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 46)
AUTHORS Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S., Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B., White,Y., Wylie,T., Waterston,R. and Wilson,R.
TITLE WashU-NCI human EST Project
JOURNAL Unpublished (1997)
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Trace considered overall poor quality
Possible reversed clone: similarity on wrong strand
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 1.
FEATURES Location/Qualifiers
source 1. .46
/organism="Homo sapiens"
/db_xref="GDB:1304390"
/db_xref="taxon:9606"
/clone="IMAGE:412696"
/clone_lib="Soares_pineal_gland_N3HPG"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: pineal gland; Vector: pT7T3D (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGGAGCGGCCGCCGTTTTTTTTTTTTTTTTTT 3'], double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT7T3 vector (Pharmacia). Library constructed by Bento Soares and M.Fatima Bonaldo. "
BASE COUNT 12 a 12 c 8 g 14 t
ORIGIN

Query Match 71.0%; Score 14.2; DB 9; Length 46;
Best Local Similarity 84.2%; Pred. No. 1.4e+04;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy      2 GGGAATAGATGAAGTTGCC 20
          || |||||||||||||| |
Db     26 GGAGATAGATGAAGTTGGC 8
```

RESULT 5
D19120
LOCUS D19120 51 bp mRNA linear EST 12-DEC-1995
DEFINITION MUSGS01332 Mouse 3'-directed Mus musculus domesticus cDNA clone mc1423 3', mRNA sequence.
ACCESSION D19120
VERSION D19120.1 GI:1089772
KEYWORDS EST.
SOURCE western European house mouse.
ORGANISM Mus musculus domesticus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 51)
AUTHORS Kawamoto,S., Okubo,K., Yoshii,J., Katsuki,M. and Matsubara,K.
TITLE Analysis of gene expression in mouse embryogenesis by 3'-directed cDNA sequencing
JOURNAL Unpublished (1995)
COMMENT Contact: Kawamoto,S., Okubo,K., Yoshii,J., Katsuki,M. and Matsubara,K.
Institute for Cellular and Molecular Biology
Osaka University
3-1 Yamada-oka, Suita, Osaka 565, Japan.
FEATURES Location/Qualifiers
source 1. .51
/organism="Mus musculus domesticus"
/strain="C57BL/6J"
/db_xref="taxon:10092"
/clone="mc1423"
/clone_lib="Mouse 3'-directed"
/tissue_type="decidual tissue (day 6.5-8.5 of gestation)"
BASE COUNT 16 a 11 c 8 g 16 t
ORIGIN

Query Match 71.0%; Score 14.2; DB 14; Length 51;
Best Local Similarity 84.2%; Pred. No. 1.4e+04;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy      1 TGGGAATAGATGAAGTTGC 19
          |||||||| || ||||| |
Db     23 TGGGAATAAATTAAGTTTC 41
```

RESULT 6
BH813380
LOCUS BH813380 95 bp DNA linear GSS 02-MAY-2002
DEFINITION SALK_064070 Arabidopsis thaliana TDNA insertion lines Arabidopsis thaliana genomic clone SALK_064070, DNA sequence.
ACCESSION BH813380
VERSION BH813380.1 GI:20391849
KEYWORDS GSS.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 95)
AUTHORS Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R., Gadrinab,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L., Shinn,P., Zimmerman,J. and Ecker,J.R.
TITLE A Sequence-Indexed Library of Insertion Mutations in the Arabidopsis Genome
JOURNAL Unpublished (2001)
COMMENT Contact: Joseph R. Ecker
Salk Institute Genomic Analysis Laboratory (SIGnAL)
The Salk Institute for Biological Studies
10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
Tel: 858 453 4100 x1752
Fax: 858 558 6379
Email: ecker@salk.edu
This is single pass sequence recovered from the left border of TDNA. This sequence lies within an annotated intron of At5g04680.
Class: TDNA tagged.

```
FEATURES             Location/Qualifiers
     source          1. .95
                     /organism="Arabidopsis thaliana"
                     /strain="Columbia 0"
                     /db_xref="taxon:3702"
                     /clone="SALK_064070"
                     /clone_lib="Arabidopsis thaliana TDNA insertion lines"
                     /note="PCR was performed on Arabidopsis thaliana lines
                     each of which contains  one or more TDNA insertion
                     elements. The resultant fragment for each  line was
                     directly sequenced to determine the genomic sequence at
                     the  site of insertion. Details of the protocols used can
                     be found at  http://signal.salk.edu/tdna_protocols.html"
BASE COUNT       33 a     13 c     22 g     27 t
ORIGIN

  Query Match             70.0%;  Score 14;  DB 17;  Length 95;
  Best Local Similarity   100.0%;  Pred. No. 2e+04;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3 GGAATAGATGAAGT 16
           ||||||||||||||
Db      43 GGAATAGATGAAGT 56


RESULT 7
AW432479
LOCUS       AW432479                 81 bp    mRNA    linear   EST 03-DEC-2001
DEFINITION  sh74f05.y1 Gm-c1015 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
            Gm-c1015-5530 5', mRNA sequence.
ACCESSION   AW432479
VERSION     AW432479.1  GI:6963786
KEYWORDS    EST.
SOURCE      soybean.
  ORGANISM  Glycine max
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
            Glycine.
REFERENCE   1  (bases 1 to 81)
  AUTHORS   Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Coryell,V., Khanna
            ,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
            Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers
            ,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk
            ,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann
            ,R., Waterston,R. and Wilson,R.
  TITLE     Public Soybean EST Project
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Shoemaker R/Public Soybean EST Project
            Public Soybean EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available through: ResGen, Invitrogen Corp. 2130
            South Memorial Parkway Huntsville, AL 35801 For further information
            call: (800)-533-4363 or contact via email: ccu@resgen.com
            Seq primer: -40RP from Gibco.
FEATURES             Location/Qualifiers
     source          1. .81
                     /organism="Glycine max"
                     /db_xref="taxon:3847"
                     /clone="GENOME SYSTEMS CLONE ID: Gm-c1015-5530"
                     /clone_lib="Gm-c1015"
                     /tissue_type="Mature flowers, field grown plants"
                     /lab_host="XL10-Gold"
                     /note="Vector: pBluescript II XR; Site_1: EcoRI; Site_2:
                     XhoI; This cDNA library was constructed from mRNA isolated
                     from mature flowers of field grown plants. The cDNA
                     library was prepared using the Stratagene pBluescript II
                     XR cDNA library construction kit. Complementary DNA was
                     synthesized from mRNA using a primer consisting of a poly
                     (dT) sequence with a XhoI restriction site. EcoRI adapters
                     were ligated to the blunt-ended cDNA fragments followed by
                     XhoI digestion. The cDNA fragments were directionally
                     cloned into the EcoRI-XhoI restriction site of the
                     pBluescript vector. The ligated cDNA fragments were
                     transformed into XL10-Gold host cells. This library was
                     constructed by Dr. Randy Shoemaker and Dr. John
                     Erpelding."
BASE COUNT       29 a      6 c     16 g     30 t
ORIGIN

  Query Match             69.0%;  Score 13.8;  DB 10;  Length 81;
  Best Local Similarity   88.2%;  Pred. No. 2.4e+04;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       4 GAATAGATGAAGTTGCC 20
           ||| |||||||||  |||
Db      18 GAAGAGATGAAGTAGCC 34


RESULT 8
AA737989
LOCUS       AA737989                 79 bp    mRNA    linear   EST 22-JAN-1998
DEFINITION  nx11f12.s1 NCI_CGAP_GC3 Homo sapiens cDNA clone IMAGE:1255823 3'
            similar to TR:Q14989 Q14989 ORF 2 ;, mRNA sequence.
ACCESSION   AA737989
VERSION     AA737989.1  GI:2768746
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 79)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html

            Trace considered overall poor quality
            Insert Length: 932   Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .79
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1255823"
                     /clone_lib="NCI_CGAP_GC3"
                     /tissue_type="pooled germ cell tumors"
                     /lab_host="DH10B"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; 1st strand cDNA was prepared from 3 pooled
                     germ cell tumors, and was then primed with a Not I -
                     oligo(dT) primer. Double-stranded cDNA was ligated to Eco
                     RI adaptors (Pharmacia), digested with Not I and cloned
                     into the Not I and Eco RI sites of the modified pT7T3
                     vector.  Library is not normalized. Library was
                     constructed by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT       29 a     11 c     16 g     23 t
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 9;  Length 79;
  Best Local Similarity   80.0%;  Pred. No. 3e+04;
```

```
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 TGGGAATAGATGAAGTTGCC 20
          ||||||| |||||  ||| |
Db     30 TGGGAATTAATGAATTTGGC 49


RESULT 9
AI680520
LOCUS       AI680520                  90 bp    mRNA    linear   EST 15-DEC-1999
DEFINITION  tw82a01.x1 NCI_CGAP_Ut3 Homo sapiens cDNA clone IMAGE:2266152 3'
            similar to gb:L25879 EPOXIDE HYDROLASE (HUMAN);contains element
            MSR1 L1 repetitive element ;, mRNA sequence.
ACCESSION   AI680520
VERSION     AI680520.1  GI:4890702
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 90)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 1400   Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .90
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2266152"
                     /clone_lib="NCI_CGAP_Ut3"
                     /tissue_type="poorly-differentiated endometrial
                     adenocarcinoma, 2 pooled tumors"
                     /lab_host="DH10B"
                     /note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.45 kb. Life Technologies catalog #:
                     11541-018"
BASE COUNT       17 a     23 c     31 g     19 t
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 9;  Length 90;
  Best Local Similarity   80.0%;  Pred. No. 3.1e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 TGGGAATAGATGAAGTTGCC 20
          |||||  |||||| | ||||
Db     43 TGGGAGGAGATGATGGTGCC 62


RESULT 10
BF465255/c
LOCUS       BF465255                  91 bp    mRNA    linear   EST 04-DEC-2000
DEFINITION  UI-M-CG0p-bqh-d-12-0-UI.s1 NIH_BMAP_Ret4_S2 Mus musculus cDNA clone
            UI-M-CG0p-bqh-d-12-0-UI 3', mRNA sequence.
ACCESSION   BF465255
VERSION     BF465255.1  GI:11534438
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 91)
  AUTHORS   Bonaldo,M.F., Lennon,G. and Soares,M.B.
  TITLE     Normalization and subtraction: two approaches to facilitate gene
            discovery
  JOURNAL   Genome Res. 6 (9), 791-806 (1996)
  MEDLINE   97044477
COMMENT     Contact: Chin, H
            National Institute of Mental Health
            6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
            20892-9643, USA
            Tel: 301 443 1706
            Fax: 301 443 9890
            Email: mEST@mail.nih.gov
            The sequence contained an oligo-dT track that was present in the
            oligonucleotide that was used to prime the synthesis of first
            strand cDNA and therefore this may represent a bonafide poly A
            tail. The sequence tag present in the cDNA between the NotI site
            and the oligo-dT track served to identify it as a clone from the
            retina tissue cDNA Library Preparation: M.B. Soares Lab Clone
            distribution: Researchers may obtain BMAP cDNA clones from RESEARCH
            GENETICS. It should be noted that Bento Soares is generating a
            small number of additional specialized non-redundant arrays of BMAP
            cDNAs whose availability will be considered under appropriate and
            limited collaborative arrangements
            Seq primer: M13 Forward
            POLYA=Yes.
FEATURES             Location/Qualifiers
     source          1. .91
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UI-M-CG0p-bqh-d-12-0-UI"
                     /clone_lib="NIH_BMAP_Ret4_S2"
                     /lab_host="DH10B (Life Technologies)"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; The
                     NIH_BMAP_Ret4_S2 library is a subtracted library,
                     ultimately derived from mouse retina tissue libraries at
                     various stages of development. For a detailed description
                     of the library from which this clone was derived, please
                     visit our web site at brainest.eng.uiowa.edu. The tissue
                     for this library was contributed by Dr. Xin-Yuan Fu, Yale
                     University School of Medicine
                     TAG_LIB=NIH_BMAP_Ret4_S2
                     TAG_TISSUE=adult-retina
                     TAG_SEQ=GTCAGCGCGCAT"
BASE COUNT       18 a     30 c     24 g     19 t
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 12;  Length 91;
  Best Local Similarity   80.0%;  Pred. No. 3.1e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 TGGGAATAGATGAAGTTGCC 20
          ||| | ||| ||||| ||||
Db     72 TGGAACTAGCTGAAGATGCC 53


RESULT 11
W80226/c
LOCUS       W80226                    93 bp    mRNA    linear   EST 25-JUN-1996
DEFINITION  me89f10.r1 Soares mouse embryo NbME13.5 14.5 Mus musculus cDNA
            clone IMAGE:402763 5' similar to gb:Z36774 M.musculus mRNA for
            alpha-2 antiplasmin (MOUSE);, mRNA sequence.
ACCESSION   W80226
VERSION     W80226.1  GI:1391209
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
```

```
REFERENCE   1  (bases 1 to 93)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:246531
            Trace considered overall poor quality
            Seq primer: ETPrimer
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .93
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:402763"
                     /clone_lib="Soares mouse embryo NbME13.5 14.5"
                     /sex="unknown"
                     /tissue_type="embryo"
                     /dev_stage="13.5-14.5dpc total fetus"
                     /lab_host="DH10B"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                     was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCGGAAATTTTTTTTTTTTTTTTTTTTTTT
                     T 3'], on equal amounts of mRNA from 2 13.5dpc and 2
                     14.5dpc embryos [total RNA provided by  Minoru Ko, Wayne
                     State Univ., from 2 ]; double-stranded cDNA was ligated to
                     Eco RI adaptors (Pharmacia), digested with Not I and
                     cloned into the Not I and Eco RI sites of the modified
                     pT7T3 vector. Library went through one round of
                     normalization, and was constructed by Bento Soares and
                     M.Fatima Bonaldo.  "
BASE COUNT       19 a     13 c     34 g     27 t
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 14;  Length 93;
  Best Local Similarity   80.0%;  Pred. No. 3.1e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 TGGGAATAGATGAAGTTGCC 20
           |||||| | |||||| | ||
Db      36 TGGGAAGAAATGAAGATTCC 17


RESULT 12
BG063371/c
LOCUS       BG063371                100 bp    mRNA    linear   EST 26-JAN-2001
DEFINITION  H3006E03-3 NIA Mouse 15K cDNA Clone Set Mus musculus cDNA clone
            H3006E03 3', mRNA sequence.
ACCESSION   BG063371
VERSION     BG063371.1  GI:12545934
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 100)
  AUTHORS   Kargul,G.J., Dudekula,D.B., Qian,Y., Lim,M.K., Jaradat,S.A., Tanaka
            ,T.S., Carter,M.G. and Ko,M.S.H.
  TITLE     Verification and initial annotation of NIA mouse 15K cDNA clone set
  JOURNAL   Unpublished (2001)
COMMENT     Other_ESTs: H3006E03-5
            Contact: George J. Kargul
            Laboratory of Genetics
            National Institute on Aging/National Institutes of Health
            333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
            Email: cdna@lgsun.grc.nia.nih.gov
            This clone set has been freely distributed to the community. Please
            visit http://lgsun.grc.nia.nih.gov/cDNA/15k.html for details.
            Plate: H3006  row: E  column: 03
            Seq primer: -21M13 Forward
            High quality sequence stop: 100
            POLYA=Yes.
FEATURES             Location/Qualifiers
     source          1. .100
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="niaEST:H3006E03-3"
                     /db_xref="taxon:10090"
                     /clone="H3006E03"
                     /clone_lib="NIA Mouse 15K cDNA Clone Set"
                     /sex="Clones arrayed from a variety of cDNA libraries"
                     /dev_stage="Clones arrayed from a variety of cDNA
                     libraries"
                     /lab_host="DH10B"
                     /note="Vector: pSPORT1; Site_1: SalI; Site_2: NotI; This
                     clone is among a rearrayed set of 15,247 clones from 11
                     embryo  cDNA libraries (including preimplantation stage
                     embryos from  unfertilized egg to blastocyst, embryonic
                     part of E7.5 embryos,  extraembryonic part of E7.5 embryos
                     , and E12.5 female mesonephros/gonad)  and one newborn
                     ovary cDNA library. Average insert size 1.5 kb. All
                     source libraries are cloned unidirectionally with Oligo(dT
                     )-Not primers.  References include: (1) Genome-wide
                     expression profiling of mid-gestation  placenta and embryo
                     using a 15,000 mouse developmental cDNA microarray,  2000,
                     Proc. Natl. Acad. Sci. U S A, 97: 9127-9132; (2)
                     Large-scale cDNA  analysis reveals phased gene expression
                     patterns during preimplantation  mouse develolpment, 2000,
                     Development, 127: 1737-1749; (3) Genome-wide  mapping of
                     unselected transcripts from extraembryonic tissue of
                     7.5-day  mouse embryos reveals enrichment in the t-complex
                     and under-representation  on the X chromosome, 1998, Hum
                     Mol Genet 7: 1967-1978."
BASE COUNT       30 a     24 c     34 g     12 t
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 12;  Length 100;
  Best Local Similarity   80.0%;  Pred. No. 3.2e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 TGGGAATAGATGAAGTTGCC 20
           |||||||| ||  |||| ||
Db      26 TGGGAATAAATTTAGTTTCC 7


RESULT 13
AI453782
LOCUS       AI453782                49 bp    mRNA    linear   EST 13-APR-1999
DEFINITION  tj23h02.x1 NCI_CGAP_Gas4 Homo sapiens cDNA clone IMAGE:2142387 3'
            similar to SW:TRAI_HUMAN P50591 TNF-RELATED APOPTOSIS INDUCING
            LIGAND ;, mRNA sequence.
ACCESSION   AI453782
VERSION     AI453782.1  GI:4284663
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 49)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
```

```
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html

            Trace considered overall poor quality
            Insert Length: 966   Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 1.
FEATURES              Location/Qualifiers
     source           1. .49
                      /organism="Homo sapiens"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:2142387"
                      /clone_lib="NCI_CGAP_Gas4"
                      /tissue_type="poorly differentiated adenocarcinoma with
                      signet ring cell features"
                      /lab_host="DH10B"
                      /note="Organ: stomach; Vector: pCMV-SPORT6; Site_1: SalI;
                      Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                      Average insert size 1.69 kb. Life Technologies catalog #:
                      11549-011"
BASE COUNT       21 a      5 c      8 g     15 t
ORIGIN

  Query Match             67.0%;  Score 13.4;  DB 9;  Length 49;
  Best Local Similarity   93.3%;  Pred. No. 3.3e+04;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 TGGGAATAGATGAAG 15
           ||||||||  ||||||
Db      33 TGGGAATAAATGAAG 47


RESULT 14
TA150D05Q/c
LOCUS       TA150D05Q                 67 bp    DNA     linear   GSS 13-DEC-2000
DEFINITION  T. brucei sheared genomic DNA clone 150d05, reverse sequence,
            genomic survey sequence.
ACCESSION   AL472813
VERSION     AL472813.1  GI:11838146
KEYWORDS    GSS.
SOURCE      Trypanosoma brucei.
  ORGANISM  Trypanosoma brucei
            Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
            Trypanosoma.
REFERENCE   1  (bases 1 to 67)
  AUTHORS   Hall,N., Bowman,S., Lennard,N.J., Doggett,J., Atkin,R.,
            Chillingworth,C., Ormond,D., Harris,B., El-Sayed,N., Hou,L.,
            Melville,S.E., Rajandream,M.A. and Barrell,B.G.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-DEC-2000) Trypanosoma brucei genome sequencing
            project, Sanger Centre, The Wellcome Trust Genome Campus, Hinxton,
            Cambridge CB10 1SA, E-mail: barrell@sanger.ac.uk and
            nh1@sanger.ac.uk
COMMENT     Constructed at the Institute for Genomic Research (TIGR),
            Rockville, MD. Genomic DNA isolated from a cloned population of
            Trypanosoma brucei (TREU927/4 GUTat 10.1) was mechanically sheared
            to give a tight size distribution (
            4 kb).  The v + i method used for the library construction is
            described in detail in Smith, H. and Venter, J.C. (Making small
            insert libraries for whole genome shotgun sequencing projects.  In
            Genome Sequencing: A Practical Approach, eds. M. Vaudin and B.
            Barrell, Oxford University Press, 1999).
            Email: nelsayed@tigr.org
            Details of T. brucei sequencing at the Sanger Centre are available
            at http://www.sanger.ac.uk/Projects/T_brucei/.
FEATURES              Location/Qualifiers
     source           1. .67
                      /organism="Trypanosoma brucei"
                      /strain="TREU927"
                      /db_xref="taxon:5691"
                      /clone="150d05"
BASE COUNT       24 a     17 c     12 g     14 t
ORIGIN

  Query Match             67.0%;  Score 13.4;  DB 17;  Length 67;
  Best Local Similarity   93.3%;  Pred. No. 3.5e+04;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       6 ATAGATGAAGTTGCC 20
           |||||||||  |||||
Db      46 ATAGATGAACTTGCC 32


RESULT 15
BI452616/c
LOCUS       BI452616                  73 bp    mRNA    linear   EST 21-AUG-2001
DEFINITION  603169810F1 NCI_CGAP_Mam5 Mus musculus cDNA clone IMAGE:5249235 5',
            mRNA sequence.
ACCESSION   BI452616
VERSION     BI452616.1  GI:15243272
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 73)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Lothar Hennighausen Ph.D., Robin Humphreys
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11629  row: d  column: 04
            High quality sequence stop: 73.
FEATURES              Location/Qualifiers
     source           1. .73
                      /organism="Mus musculus"
                      /strain="C57BL/6J"
                      /db_xref="taxon:10090"
                      /clone="IMAGE:5249235"
                      /clone_lib="NCI_CGAP_Mam5"
                      /tissue_type="tumor, gross tissue"
                      /dev_stage="7 months"
                      /lab_host="DH10B"
                      /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
                      Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                      Library constructed by Life Technologies. Investigators
                      providing samples: Lothar Hennighausen/Robin Humphreys,
                      NIH"
BASE COUNT       29 a     15 c     13 g     16 t
ORIGIN

  Query Match             67.0%;  Score 13.4;  DB 13;  Length 73;
  Best Local Similarity   93.3%;  Pred. No. 3.6e+04;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       6 ATAGATGAAGTTGCC 20
           |||||||||||||  |
Db      17 ATAGATGAAGTTGTC 3


Search completed: November 26, 2002, 04:08:46
```

Job time : 765.8 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 3, 2002, 14:46:40 ; Search time 302.2 Seconds
(without alignments)
1926.063 Million cell updates/sec

Title: US-09-296-264-16
Perfect score: 20
Sequence: 1 tcctctggcttctggtagcg 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 995600

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
c	1	15.4	77.0	50	6	AR032584	AR032584 Sequence
c	2	15.4	77.0	50	6	AR209248	AR209248 Sequence
c	3	15.4	77.0	50	6	I29324	I29324 Sequence 19
c	4	15.4	77.0	50	6	I90998	I90998 Sequence 19
	5	13.8	69.0	28	6	E34315	E34315 Method for
	6	13.8	69.0	100	6	AX328146	AX328146 Sequence
	7	13.6	68.0	31	6	A84905	A84905 Sequence 54
	8	13.6	68.0	51	6	AX105423	AX105423 Sequence
c	9	13.6	68.0	51	6	AX105424	AX105424 Sequence
	10	13.6	68.0	51	6	AX417936	AX417936 Sequence
	11	13.6	68.0	65	6	AX483733	AX483733 Sequence
	12	13.4	67.0	51	6	AX199175	AX199175 Sequence
c	13	13.4	67.0	54	6	AR040512	AR040512 Sequence
c	14	13.2	66.0	20	6	AX104210	AX104210 Sequence
c	15	13.2	66.0	20	6	AX355315	AX355315 Sequence
	16	13.2	66.0	51	10	U92134	U92134 Mus musculu
	17	13.2	66.0	73	6	AR035230	AR035230 Sequence
c	18	13.2	66.0	77	6	AR035229	AR035229 Sequence
c	19	13.2	66.0	77	6	AR035231	AR035231 Sequence
	20	13.2	66.0	99	3	AB002396	AB002396 Bombyx mo
c	21	13	65.0	27	6	AX116180	AX116180 Sequence
	22	13	65.0	51	6	AX117049	AX117049 Sequence
c	23	12.8	64.0	22	6	AR108112	AR108112 Sequence
	24	12.8	64.0	27	6	AX350717	AX350717 Sequence
c	25	12.8	64.0	47	6	AX147932	AX147932 Sequence
c	26	12.8	64.0	48	6	AX147903	AX147903 Sequence
	27	12.8	64.0	60	6	AR006759	AR006759 Sequence
	28	12.8	64.0	60	6	I71271	I71271 Sequence 9
c	29	12.8	64.0	95	9	HUMDCCAB	M63700 Human tumor
	30	12.8	64.0	98	9	HSU08253	U08253 Human chrom
c	31	12.6	63.0	21	6	AX166288	AX166288 Sequence
c	32	12.6	63.0	30	6	AX127355	AX127355 Sequence
	33	12.6	63.0	30	6	AX473006	AX473006 Sequence
c	34	12.6	63.0	32	6	AR099474	AR099474 Sequence
c	35	12.6	63.0	32	6	AR178755	AR178755 Sequence
c	36	12.6	63.0	35	6	AR074031	AR074031 Sequence
c	37	12.6	63.0	35	6	I50888	I50888 Sequence 17
	38	12.6	63.0	39	6	AR099119	AR099119 Sequence
	39	12.6	63.0	42	10	MMU232720	AJ232720 Mus muscu
	40	12.6	63.0	42	10	MMU232721	AJ232721 Mus muscu
c	41	12.6	63.0	50	10	MMTFAM11	U71388 Mus musculu
	42	12.6	63.0	51	6	E14329	E14329 Primer. 7/1
	43	12.6	63.0	56	6	AR061920	AR061920 Sequence
	44	12.6	63.0	56	6	E06683	E06683 DNA probe t
c	45	12.6	63.0	68	6	I34324	I34324 Sequence 23

ALIGNMENTS

RESULT 1
AR032584/c
LOCUS AR032584 50 bp DNA linear PAT 29-SEP-1999
DEFINITION Sequence 196 from patent US 5869241.
ACCESSION AR032584
VERSION AR032584.1 GI:5948189
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 50)
AUTHORS Edwards,C.A., Cantor,C.R., Andrews,B.M., Turin,L.M. and Fry,K.E.
TITLE Method of determining DNA sequence preference of a DNA-binding molecule
JOURNAL Patent: US 5869241-A 196 09-FEB-1999;

```
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="unknown"
BASE COUNT       16 a     11 c     13 g     10 t
ORIGIN

  Query Match             77.0%;  Score 15.4;  DB 6;  Length 50;
  Best Local Similarity   94.1%;  Pred. No. 4.5e+03;
  Matches   16;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 TCCTCTGGCTTCTGGTA 17
          |||||||||| ||||||
Db     45 TCCTCTGGCTGCTGGTA 29


RESULT 2
AR209248/c
LOCUS       AR209248                 50 bp    DNA     linear   PAT 20-JUN-2002
DEFINITION  Sequence 196 from patent US 6384208.
ACCESSION   AR209248
VERSION     AR209248.1  GI:21510616
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Edwards,C.A., Cantor,C.R., Andrews,B.M., Turin,L.M. and Fry,K.E.
  TITLE     Sequence directed DNA binding molecules compositions and methods
  JOURNAL   Patent: US 6384208-A 196 07-MAY-2002;
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="unknown"
BASE COUNT       16 a     11 c     13 g     10 t
ORIGIN

  Query Match             77.0%;  Score 15.4;  DB 6;  Length 50;
  Best Local Similarity   94.1%;  Pred. No. 4.5e+03;
  Matches   16;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 TCCTCTGGCTTCTGGTA 17
          |||||||||| ||||||
Db     45 TCCTCTGGCTGCTGGTA 29


RESULT 3
I29324/c
LOCUS       I29324                   50 bp    DNA     linear   PAT 06-FEB-1997
DEFINITION  Sequence 196 from patent US 5578444.
ACCESSION   I29324
VERSION     I29324.1  GI:1820115
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Edwards,C.A., Cantor,C.R., Andrews,B.M., Turin,L.M. and Fry,K.E.
  TITLE     Sequence-directed DNA-binding molecules compositions and methods
  JOURNAL   Patent: US 5578444-A 196 26-NOV-1996;
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="unknown"
BASE COUNT       16 a     11 c     13 g     10 t
ORIGIN

  Query Match             77.0%;  Score 15.4;  DB 6;  Length 50;
  Best Local Similarity   94.1%;  Pred. No. 4.5e+03;
  Matches   16;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 TCCTCTGGCTTCTGGTA 17
          |||||||||| ||||||
Db     45 TCCTCTGGCTGCTGGTA 29


RESULT 4
I90998/c
LOCUS       I90998                   50 bp    DNA     linear   PAT 01-DEC-1998
DEFINITION  Sequence 196 from patent US 5726014.
ACCESSION   I90998
VERSION     I90998.1  GI:3935468
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Edwards,C.A., Cantor,C.R., Andrews,B.M. and Turin,L.M.
  TITLE     Screening assay for the detection of DNA-binding molecules
  JOURNAL   Patent: US 5726014-A 196 10-MAR-1998;
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="unknown"
BASE COUNT       16 a     11 c     13 g     10 t
ORIGIN

  Query Match             77.0%;  Score 15.4;  DB 6;  Length 50;
  Best Local Similarity   94.1%;  Pred. No. 4.5e+03;
  Matches   16;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 TCCTCTGGCTTCTGGTA 17
          |||||||||| ||||||
Db     45 TCCTCTGGCTGCTGGTA 29


RESULT 5
E34315
LOCUS       E34315                   28 bp    DNA     linear   PAT 31-JAN-2002
DEFINITION  Method for detecting gene deletion.
ACCESSION   E34315
VERSION     E34315.1  GI:18624300
KEYWORDS    JP 2000093185-A/3.
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 28)
  AUTHORS   Kanno,Y., Nomura,Y. and Yamaguchi,T.
  TITLE     Method for detecting gene deletion
  JOURNAL   Patent: JP 2000093185-A 3 04-APR-2000;
            JAPAN FOUND CANCER,BML INC
COMMENT     OS   Unidentified
            PN   JP 2000093185-A/3
            PD   04-APR-2000
            PF   25-SEP-1998 JP 1998288796
            PR
            PI   YASUYOSHI KANNO,YUKIO NOMURA,TOSHIKAZU YAMAGUCHI PC
            C12N15/09,C12Q1/68,C12N15/00
            CC   Strandedness: Single;
            CC   Topology: Linear;
            FH   Key             Location/Qualifiers
            FT   source          1. .28
            FT                   /organism='Unidentified'.
FEATURES             Location/Qualifiers
     source          1. .28
                     /organism="unidentified"
                     /db_xref="taxon:32644"
BASE COUNT        6 a      7 c      4 g     11 t
ORIGIN

  Query Match             69.0%;  Score 13.8;  DB 6;  Length 28;
  Best Local Similarity   88.2%;  Pred. No. 3.2e+04;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      3 CTCTGGCTTCTGGTAGC 19
          |||||||||||  ||||
Db      1 CTCTGGCTTCTTTTAGC 17


RESULT 6
```

```
AX328146
LOCUS       AX328146                100 bp    DNA     linear   PAT 07-JAN-2002
DEFINITION  Sequence 22 from Patent WO0189559.
ACCESSION   AX328146
VERSION     AX328146.1  GI:18098189
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Audonnet,J.C., Bublot,M.J., Perez,J.M. and Baudu,P.G.
  TITLE     Porcine reproductive and respiratory syndrome virus (prrsv)
            recombinant poxvirus vaccine
  JOURNAL   Patent: WO 0189559-A 22 29-NOV-2001;
            MERIAL (FR)
FEATURES             Location/Qualifiers
     source          1. .100
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="oligonucleotide"
BASE COUNT       17 a     20 c     23 g     40 t
ORIGIN

  Query Match             69.0%;  Score 13.8;  DB 6;  Length 100;
  Best Local Similarity   88.2%;  Pred. No. 2.9e+04;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      3 CTCTGGCTTCTGGTAGC 19
          |||| ||||||||| ||
Db     71 CTCTTGCTTCTGGTGGC 87


RESULT 7
A84905
LOCUS       A84905                   31 bp    DNA     linear   PAT 21-JAN-2000
DEFINITION  Sequence 54 from Patent WO9844106.
ACCESSION   A84905
VERSION     A84905.1  GI:6733753
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 31)
  AUTHORS   Waeber,G. and Bonny,C.
  TITLE     TRANSCRIPTION FACTOR ISLET-BRAIN 1 (IB1)
  JOURNAL   Patent: WO 9844106-A 54 08-OCT-1998;
            WAEBER GERARD (CH); NICOD PASCAL (CH)
FEATURES             Location/Qualifiers
     source          1. .31
                     /organism="unidentified"
                     /db_xref="taxon:32644"
BASE COUNT        4 a      8 c      9 g     10 t
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 6;  Length 31;
  Best Local Similarity   80.0%;  Pred. No. 4.1e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 TCCTCTGGCTTCTGGTAGCG 20
          ||| ||||||||| | || |
Db      8 TCCCCTGGCTTCTAGGAGAG 27


RESULT 8
AX105423
LOCUS       AX105423                 51 bp    DNA     linear   PAT 30-APR-2001
DEFINITION  Sequence 41 from Patent WO0123533.
ACCESSION   AX105423
VERSION     AX105423.1  GI:13921532
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 51)
  AUTHORS   Gurney,M. and Bienkowski,M.J.
  TITLE     Alzheimer's disease secretase, app substrates therefor, and uses
            therefor
  JOURNAL   Patent: WO 0123533-A 41 05-APR-2001;
            Pharmacia & Upjohn Company (US)
FEATURES             Location/Qualifiers
     source          1. .51
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="Caspase 8 Cleavage Site"
BASE COUNT       11 a     15 c     12 g     13 t
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 6;  Length 51;
  Best Local Similarity   80.0%;  Pred. No. 3.9e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 TCCTCTGGCTTCTGGTAGCG 20
          ||| ||||  ||||||| ||
Db     23 TCCGCTGGACTCTGGTATCG 42


RESULT 9
AX105424/c
LOCUS       AX105424                 51 bp    DNA     linear   PAT 30-APR-2001
DEFINITION  Sequence 42 from Patent WO0123533.
ACCESSION   AX105424
VERSION     AX105424.1  GI:13921533
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 51)
  AUTHORS   Gurney,M. and Bienkowski,M.J.
  TITLE     Alzheimer's disease secretase, app substrates therefor, and uses
            therefor
  JOURNAL   Patent: WO 0123533-A 42 05-APR-2001;
            Pharmacia & Upjohn Company (US)
FEATURES             Location/Qualifiers
     source          1. .51
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="Caspase 8 Cleavage Site"
BASE COUNT       13 a     12 c     15 g     11 t
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 6;  Length 51;
  Best Local Similarity   80.0%;  Pred. No. 3.9e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 TCCTCTGGCTTCTGGTAGCG 20
          ||| ||||  ||||||| ||
Db     33 TCCGCTGGACTCTGGTATCG 14


RESULT 10
AX417936
LOCUS       AX417936                 51 bp    DNA     linear   PAT 18-JUN-2002
DEFINITION  Sequence 43 from Patent WO0230945.
ACCESSION   AX417936
VERSION     AX417936.1  GI:21523048
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Lasters,I., Jespers,L., Wang,P. and Winter,G.
  TITLE     Concatenated nucleic acid sequences
  JOURNAL   Patent: WO 0230945-A 43 18-APR-2002;
            MEDICAL RES COUNCIL (GB)
FEATURES             Location/Qualifiers
     source          1. .51
```

```
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="synthetic onigonucleotide (L J 805"
BASE COUNT        6 a      12 c      19 g      14 t
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 6;  Length 51;
  Best Local Similarity   80.0%;  Pred. No. 3.9e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 TCCTCTGGCTTCTGGTAGCG 20
          |||||||||||  || | | |
Db     18 TCCTCTGGCTTGTGCTGGTG 37


RESULT 11
AX483733
LOCUS       AX483733                 65 bp    DNA     linear   PAT 16-AUG-2002
DEFINITION  Sequence 1033 from Patent WO02053728.
ACCESSION   AX483733
VERSION     AX483733.1  GI:22318085
KEYWORDS    .
SOURCE      Candida albicans.
  ORGANISM  Candida albicans
            Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
            Saccharomycetales; mitosporic Saccharomycetales; Candida.
REFERENCE   1
  AUTHORS   Roemer,T., Jiang,B., Boone,C., Bussey,H. and Ohlsen,K.L.
  TITLE     Gene disruption methodologies for drug target discovery
  JOURNAL   Patent: WO 02053728-A 1033 11-JUL-2002;
            Elitra Pharmaceuticals, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .65
                     /organism="Candida albicans"
                     /db_xref="taxon:5476"
BASE COUNT       17 a     17 c      7 g     24 t
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 6;  Length 65;
  Best Local Similarity   80.0%;  Pred. No. 3.8e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 TCCTCTGGCTTCTGGTAGCG 20
          ||||||  |||||  ||| ||
Db     41 TCCTCTATCTTCTAGTATCG 60


RESULT 12
AX199175
LOCUS       AX199175                 51 bp    DNA     linear   PAT 29-AUG-2001
DEFINITION  Sequence 105 from Patent WO0151670.
ACCESSION   AX199175
VERSION     AX199175.1  GI:15389527
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 51)
  AUTHORS   Shimkets,R.A. and Leach,M.D.
  TITLE     Nucleic acids containing single nucleotide polymorphisms and
            methods of use thereof
  JOURNAL   Patent: WO 0151670-A 105 19-JUL-2001;
            Curagen Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .51
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     misc_feature    26
                     /note="1 of 2 allelic variants (106 is other entry)
                     Accession number cg43976973"
BASE COUNT        9 a     13 c      14 g     15 t
ORIGIN

  Query Match             67.0%;  Score 13.4;  DB 6;  Length 51;
  Best Local Similarity   93.3%;  Pred. No. 5e+04;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      2 CCTCTGGCTTCTGGT 16
          | |||||||||||||
Db     12 CATCTGGCTTCTGGT 26


RESULT 13
AR040512/c
LOCUS       AR040512                 54 bp    DNA     linear   PAT 29-SEP-1999
DEFINITION  Sequence 1360 from patent US 5807743.
ACCESSION   AR040512
VERSION     AR040512.1  GI:5959875
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 54)
  AUTHORS   Stinchcomb,D.T. and McSwiggen,J.A.
  TITLE     Interleukin-2 receptor gamma-chain ribozymes
  JOURNAL   Patent: US 5807743-A 1360 15-SEP-1998;
FEATURES             Location/Qualifiers
     source          1. .54
                     /organism="unknown"
BASE COUNT       18 a     15 c      12 g      9 t
ORIGIN

  Query Match             67.0%;  Score 13.4;  DB 6;  Length 54;
  Best Local Similarity   93.3%;  Pred. No. 4.9e+04;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      5 CTGGCTTCTGGTAGC 19
          |||||||||||| ||
Db     17 CTGGCTTCTGGTTGC 3


RESULT 14
AX104210/c
LOCUS       AX104210                 20 bp    DNA     linear   PAT 30-APR-2001
DEFINITION  Sequence 402 from Patent WO0122972.
ACCESSION   AX104210
VERSION     AX104210.1  GI:13920407
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 20)
  AUTHORS   Krieg,A.M., Schetter,C. and Vollmer,J.C.
  TITLE     Immunostimulatory nucleic acids
  JOURNAL   Patent: WO 0122972-A 402 05-APR-2001;
            UNIVERSITY OF IOWA RESEARCH FOUNDATION (US) ; Coley Pharmaceutical
            GmbH (DE)
FEATURES             Location/Qualifiers
     source          1. .20
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
BASE COUNT        3 a      6 c      7 g      4 t
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 6;  Length 20;
  Best Local Similarity   83.3%;  Pred. No. 6.8e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      3 CTCTGGCTTCTGGTAGCG 20
          ||| |||| | |||||||
Db     19 CTCAGGCTACCGGTAGCG 2


RESULT 15
AX355315/c
```

```
LOCUS       AX355315                 20 bp    DNA     linear   PAT 06-FEB-2002
DEFINITION  Sequence 343 from Patent WO0197843.
ACCESSION   AX355315
VERSION     AX355315.1  GI:18619983
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Weiner,G. and Hartmann,G.
  TITLE     Methods for enhancing antibody-induced cell lysis and treating
            cancer
  JOURNAL   Patent: WO 0197843-A 343 27-DEC-2001;
            UNIVERSITY OF IOWA RESEARCH FOUNDATION (US)
FEATURES             Location/Qualifiers
     source          1. .20
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="Synthetic oligonucleotide"
BASE COUNT        3 a      6 c      7 g      4 t
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 6;  Length 20;
  Best Local Similarity   83.3%;  Pred. No. 6.8e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       3 CTCTGGCTTCTGGTAGCG 20
           ||| |||| | |||||||
Db      19 CTCAGGCTACCGGTAGCG 2


Search completed: December  3, 2002, 18:13:46
Job time : 310.2 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 05:52:31 ; Search time 98.55 Seconds
(without alignments)
457.027 Million cell updates/sec

Title: US-09-296-264-16
Perfect score: 20
Sequence: 1 tcctctggcttctggtagcg 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 2390332

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		20	100.0	20	21	AAZ31446	Human neuropilin m
c 2		15.4	77.0	50	15	AAQ69446	Human immune inter
c 3		15.4	77.0	50	18	AAT63908	Human immune inter
c 4		15.4	77.0	50	20	AAX17196	Test sequence from
c 5		15.4	77.0	50	24	ABK82687	DNA binding molecu
6		15.2	76.0	60	24	ABN45339	Human spliced tran
7		14.8	74.0	60	24	ABN39404	Human spliced tran
8		14.2	71.0	23	19	AAV51908	Zea mays genome re
c 9		14.2	71.0	41	19	AAV50992	Maize polymorphic
c 10		14.2	71.0	41	19	AAV50993	Maize polymorphic
c 11		14.2	71.0	41	19	AAV50969	Maize polymorphic
c 12		14.2	71.0	41	19	AAV50971	Maize polymorphic
c 13		14.2	71.0	41	19	AAV50980	Maize polymorphic
c 14		14.2	71.0	41	19	AAV50981	Maize polymorphic
c 15		14.2	71.0	41	19	AAV47809	Maize polymorphic
c 16		14.2	71.0	41	19	AAV47810	Maize polymorphic
c 17		14.2	71.0	41	19	AAV47797	Maize polymorphic
c 18		14.2	71.0	41	19	AAV47798	Maize polymorphic
c 19		14.2	71.0	41	19	AAV47786	Maize polymorphic
c 20		14.2	71.0	41	19	AAV47788	Maize polymorphic
21		14.2	71.0	51	22	AAL26929	Human SNP oligonuc
22		14.2	71.0	51	24	AAD37790	Human STIM2 cDNA f
23		14.2	71.0	65	24	ABN53981	Mouse spliced tran
24		13.8	69.0	28	21	AAA38101	PCR primer used in
25		13.8	69.0	51	22	AAL30729	Human SNP oligonuc
c 26		13.8	69.0	92	24	ABN88559	Coagulation factor
c 27		13.8	69.0	96	24	ABN88501	Coagulation factor
28		13.8	69.0	100	24	ABK15619	PRRSV ORF 5 PCR pr
29		13.6	68.0	31	19	AAV45577	Human IB1 gene int
c 30		13.6	68.0	43	9	AAN81575	Linker DNA contg.
31		13.6	68.0	51	21	AAA15683	Caspase 8 cleavage
c 32		13.6	68.0	51	21	AAA15684	Caspase 8 cleavage
33		13.6	68.0	51	22	AAD17886	Oligo #571 used fo
c 34		13.6	68.0	51	22	AAD17887	Oligo #572 used fo
35		13.6	68.0	51	22	AAD13267	N-terminal human a
c 36		13.6	68.0	51	22	AAD13268	N-terminal human a
37		13.6	68.0	51	22	AAS11538	Sense oligonucleot
c 38		13.6	68.0	51	22	AAS11539	Antisense oligo fo
39		13.6	68.0	51	22	AAS11723	T7-caspase-caspase
c 40		13.6	68.0	51	22	AAS11724	T7-caspase-caspase
41		13.6	68.0	51	22	AAD06760	Hu-Asp2 expression
c 42		13.6	68.0	51	22	AAD06761	Hu-Asp2 expression
43		13.6	68.0	51	24	ABL52478	Human caspase 8 ol
c 44		13.6	68.0	51	24	ABL52479	Human caspase 8 ol
c 45		13.6	68.0	60	24	ABN45979	Human spliced tran

ALIGNMENTS

RESULT 1
AAZ31446
ID AAZ31446 standard; DNA; 20 BP.
XX
AC AAZ31446;
XX
DT 07-FEB-2000 (first entry)
XX
DE Human neuropilin mRNA specific antisense oligo GTI3617.
XX
KW Neuropilin; human; growth; metastasis; tumor; neovascularisation;
KW cancer; papilloma; diabetic retinopathy; antisense; ss.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9955855-A2.
XX
PD 04-NOV-1999.
XX
PF 23-APR-1999; 99WO-CA00324.
XX
PR 23-APR-1998; 98US-0082791.
XX
PA (GENE-) GENESENSE TECHNOLOGIES INC.
XX
PI Wright JA, Young AH, Lee YS;
XX
DR WPI; 2000-023357/02.
XX
PT Antisense oligonucleotides that inhibit neuropilin expression, useful
PT for treating cancer -

XX
PS Claim 4; Page 16; 57pp; English.
XX
CC Sequences AAZ31431-460 represent antisense oligonucleotides which
CC inhibit human neuropilin expression. The antisense oligonucleotides can
CC be used to inhibit the growth or metastasis of a mammalian tumor and
CC inhibit neovascularisation. The oligonucleotides may be used to treat
CC various forms of cancers or tumors, such as sarcomas, melanomas,
CC adenomas, carcinomas of solid tissue, hypoxic tumors, squamous cell
CC carcinomas of the mouth, throat, larynx and lung, genitourinary cancers
CC such as cervical and bladder cancer, hematopoietic cancers, colon cancer,
CC breast cancer, pancreatic cancer, renal cancer, brain cancer, skin
CC cancer, liver cancer, head and neck cancers, and nervous system cancers,
CC as well as benign lesions such as papillomas. The methods may be used to
CC treat neovascularisation disorders such as diabetic retinopathy, and
CC retinopathy of prematurity and age related macular degeneration.
XX
SQ Sequence 20 BP; 1 A; 6 C; 6 G; 7 T; 0 other;

Query Match 100.0%; Score 20; DB 21; Length 20;
Best Local Similarity 100.0%; Pred. No. 6.2;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCCTCTGGCTTCTGGTAGCG 20
||||||||||||||||||||
Db 1 TCCTCTGGCTTCTGGTAGCG 20

RESULT 2
AAQ69446/c
ID AAQ69446 standard; DNA; 50 BP.
XX
AC AAQ69446;
XX
DT 27-FEB-1995 (first entry)
XX
DE Human immune interferon (IFN-gamma) gene, target region.
XX
KW DNA protein-binding assay; test sequence; screening sequence;
KW promoter; target; TATA box; Herpes Simplex Virus; HSV;
KW origin of replication; UL9; transcription factor; TFIID: ds.
XX
OS Synthetic.
XX
PN WO9414980-A.
XX
PD 07-JUL-1994.
XX
PF 20-DEC-1993; 93WO-US12388.
XX
PR 23-DEC-1992; 92US-0996783.
PR 17-SEP-1993; 93US-0123936.
XX
PA (GENE-) GENELABS TECHNOLOGIES INC.
XX
PI Andrews BM, Cantor CR, Edwards CA, Fry KE, Turin LM;
XX
DR WPI; 1994-234711/28.
XX
PT Sequence-directed DNA-binding molecules - useful in
PT pharmaceuticals and as molecular reagents
XX
PS Claim 28; Page 310; 587pp; English.
XX
CC A DNA protein-binding assay is provided, useful for screening
CC libraries of synthetic or biological cpds. for their ability
CC to bind DNA test sequences. The assay is versatile in that any
CC number of test sequences can be tested by placing the test sequence
CC adjacent to a defined protein-binding screening sequence. Binding
CC of mols. to these test sequences changes the binding characteristics
CC of the protein mol. to its cognate binding sequence. When such a mol.
CC binds the test sequence, the equilibrium of the DNA:protein complexes
CC is disturbed, generating changes in the concentration of free DNA probe.
CC One application of this method is to eucaryotic general transcription
CC factors (e.g. TFIID), where the target region is typically selected
CC from DNA sequences adjacent to the binding site for the eucaryotic
CC transcription factor. Numerous exemplary test sequences are given:
CC the sequences in AAQ69251-731 and AAQ69850 correspond to promoter
CC targets (typically, TATA box-contg. sites) for human genes and the
CC sequences in AAQ69732-849 correspond to promoter targets for viral genes.
CC The test sequences may also be randomly generated. DNA:protein
CC interaction may be used for screening purposes, e.g. the Herpes Simplex
CC Virus (HSV) origin of replication and UL9 (see AAQ69851-52, AAQ69865 and
CC AAQ69891).
XX
SQ Sequence 50 BP; 16 A; 11 C; 13 G; 10 T; 0 other;

Query Match 77.0%; Score 15.4; DB 15; Length 50;
Best Local Similarity 94.1%; Pred. No. 8.4e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TCCTCTGGCTTCTGGTA 17
||||||||||.||||||
Db 45 TCCTCTGGCTGCTGGTA 29

RESULT 3
AAT63908/c
ID AAT63908 standard; DNA; 50 BP.
XX
AC AAT63908;
XX
DT 14-MAR-1997 (first entry)
XX
DE Human immune interferon (IFN-gamma) gene TFIID binding site.
XX
KW Duplex DNA; target region; binding characteristic; DNA binding protein;
KW TFIID; transcription factor; binding site; inhibition; enhance;
KW cancer; inherited genetic disorder; alpha-D-galactosidase A; ds.
XX
OS Homo sapiens.
XX
PN US5578444-A.
XX
PD 26-NOV-1996.
XX
PF 27-JUN-1991; 91US-0723618.
XX
PR 20-DEC-1993; 93US-0171389.
PR 27-JUN-1991; 91US-0723618.
PR 23-DEC-1992; 92US-0996783.
PR 17-SEP-1993; 93US-0123936.
XX
PA (GENE-) GENELABS TECHNOLOGIES INC.
XX
PI Andrews BM, Cantor CR, Edwards CA, Fry KE, Turin LM;
XX
DR WPI; 1997-020402/02.
XX
PT Altering binding characteristics of DNA binding proteins to duplex
PT DNA - by attaching specific small cpd. to target region close to the
PT protein's binding site, useful in treatment of viral disease, cancer
PT etc
XX
PS Claim 6; Column 199-200; 264pp; English.
XX
CC The sequences given in AAT63713-4312 represent duplex DNA's which act
CC as target regions in the method of the invention. The method for
CC altering the binding characteristics of a DNA-binding protein to duplex
CC DNA comprises contacting the duplex DNA with a small molecule which
CC binds sequence-specifically to a target region, where, when the small
CC molecule is bound to the target region, it is adjacent to, but not
CC overlapping by more than 4 bp, a binding site for a DNA-binding protein.
CC The small molecule is added at a concentration effective to alter the
CC binding of the DNA binding protein, pref. TFIID, to its binding site on
CC the duplex DNA. The binding of the small molecule may inhibit or

CC enhance the binding of the DNA-binding protein to its binding site. The
CC compounds isolated using this method are potentially useful as
CC therapeutic agents for treatment of any disease which involves a
CC specific DNA sequence, e.g. cancer, or inherited genetic disorders etc.
CC The method is suitable for screening large biological or chemical
CC libraries and allows determination of sequence-specific and relative
CC affinities of known DNA-binding agents for different DNA sequences.
CC The design of these duplex DNA's allows a single DNA:protein interaction
CC to be used for screening sequence-specific, or preferential, DNA binding
CC proteins that recognise almost any possible sequence (see also AAT49539-
CC 74).
XX
SQ Sequence 50 BP; 16 A; 11 C; 13 G; 10 T; 0 other;

Query Match 77.0%; Score 15.4; DB 18; Length 50;
Best Local Similarity 94.1%; Pred. No. 8.4e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TCCTCTGGCTTCTGGTA 17
||||||||||| ||||||
Db 45 TCCTCTGGCTGCTGGTA 29

RESULT 4
AAX17196/c
ID AAX17196 standard; DNA; 50 BP.
XX
AC AAX17196;
XX
DT 06-MAY-1999 (first entry)
XX
DE Test sequence from human immune interferon (INF-gamma) gene.
XX
KW Test sequence; DNA-binding molecule; screening sequence; human;
KW nucleic acid amplification; target; viral; ds.
XX
OS Homo sapiens.
XX
PN US5869241-A.
XX
PD 09-FEB-1999.
XX
PF 07-JUN-1995; 95US-0475228.
XX
PR 20-DEC-1993; 93US-0171389.
PR 27-JUN-1991; 91US-0723618.
PR 23-DEC-1992; 92US-0996783.
PR 17-SEP-1993; 93US-0123936.
PR 07-JUN-1995; 95US-0475228.
XX
PA (GENE-) GENELABS TECHNOLOGIES INC.
XX
PI Andrews BM, Cantor CR, Edwards CA, Fry KE, Turin LM;
XX
DR WPI; 1999-152755/13.
XX
PT Determination of DNA sequence preference of a DNA-binding molecule -
PT based on inhibition of binding of protein to oligonucleotide
PT sequence attached to test sequence
XX
PS Claim 3; Columns 199-200; 270pp; English.
XX
CC Sequences AAX17001 to AAX17600 represent specifically claimed target
CC test sequences that are used in the method of the invention of
CC determining the DNA sequence preference of a DNA-binding molecule. The
CC method comprises: (i) adding a test molecule and a DNA-binding protein to
CC a mixture of duplex DNA test oligonucleotides, each of the test
CC oligonucleotides having a test sequence adjacent to a screening sequence,
CC where the screening sequence binds to the DNA-binding protein with a
CC binding affinity that is independent of the DNA sequence of the test
CC sequence, and where the mixture of duplex DNA test oligonucleotides
CC includes several test sequences; (ii) incubating the test molecule, the
CC mixture of duplex DNA test oligonucleotides and the DNA-binding protein
CC for a time sufficient to permit binding of the test molecule to test
CC sequences in the duplex DNA; (iii) separating unbound test
CC oligonucleotides from test oligonucleotides bound to binding protein;
CC (iv) amplifying the unbound test oligonucleotides; (v) repeating steps
CC (ii) to (iv); (vi) isolating the amplified test oligonucleotides; and
CC (vii) sequencing the isolated test oligonucleotides. Test sequences
CC AAX17001-X17481 and AAX17600 correspond to promoter targets for human
CC genes and test sequences AAX17482-X17599 correspond to promoter targets
CC for viral genes.
XX
SQ Sequence 50 BP; 16 A; 11 C; 13 G; 10 T; 0 other;

Query Match 77.0%; Score 15.4; DB 20; Length 50;
Best Local Similarity 94.1%; Pred. No. 8.4e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TCCTCTGGCTTCTGGTA 17
||||||||||| ||||||
Db 45 TCCTCTGGCTGCTGGTA 29

RESULT 5
ABK82687/c
ID ABK82687 standard; DNA; 50 BP.
XX
AC ABK82687;
XX
DT 27-AUG-2002 (first entry)
XX
DE DNA binding molecule screening method test sequence #196.
XX
KW DNA binding molecule screening; inhibition of transcription;
KW infection; human immunodeficiency virus; HIV; parasite; cancer;
KW cardiovascular; respiratory; gastrointestinal; endocrine; metabolic;
KW rheumatic; immunological; haematological; neurological;
KW psychiatric; dermatological; ophthalmological; musculo-skeletal;
KW urogenital disorder; ss.
XX
OS Synthetic.
XX
PN US6384208-B1.
XX
PD 07-MAY-2002.
XX
PF 15-JUL-1999; 99US-0354947.
XX
PR 20-DEC-1993; 93US-0171389.
PR 07-JUN-1995; 95US-0482080.
PR 27-JUN-1991; 91US-0723618.
PR 23-DEC-1992; 92US-0996783.
PR 17-SEP-1993; 93US-0123936.
XX
PA (GENE-) GENELABS TECHNOLOGIES INC.
XX
PI Edwards CA, Cantor CR, Andrews BM, Turin LM, Fry KE;
XX
DR WPI; 2002-442819/47.
XX
PT Decreasing transcriptional activity of genes for treating infections or
PT cancer, by administration of an agent that binds to two non-overlapping
PT regions of the gene -
XX
PS Example 15; SEQ ID No 196; 98pp; English.
XX
CC The invention relates to a method of decreasing transcriptional activity
CC in a duplex deoxyribonucleic acid (DNA) template (T1) comprising
CC contacting (T1) with a binding agent comprising at least one small duplex
CC DNA-binding molecule (T2) coupled to at least one other small duplex-
CC binding molecule that binds to a non-overlapping region of target
CC sequence (TS). The method is useful for inhibiting transcription of a
CC range of disease-related genes for treating infections (by viruses,
CC including human immunodeficiency virus, bacteria, fungi, protozoa
CC and parasites), cancer, cardiovascular, respiratory, gastrointestinal,

```
CC   endocrine/metabolic, rheumatic/immunological, haematological,
CC   neurological, psychiatric, dermatological, ophthalmological,
CC   musculo-skeletal, genetic or urogenital disorders. The method provides
CC   sequence-specific inhibition of transcription of pathological genes
CC   without affecting transcription of cellular genes regulated by the same
CC   transcription factor, and can be applied to regulation of any gene.
CC   ABK82492-ABK83155 represent DNA binding molecule test sequences used in
CC   the method of the invention.
XX
SQ   Sequence 50 BP; 16 A; 11 C; 13 G; 10 T; 0 other;

  Query Match             77.0%;  Score 15.4;  DB 24;  Length 50;
  Best Local Similarity   94.1%;  Pred. No. 8.4e+02;
  Matches   16;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 TCCTCTGGCTTCTGGTA 17
           |||||||||| ||||||
Db      45 TCCTCTGGCTGCTGGTA 29


RESULT 6
ABN45339
ID   ABN45339 standard; DNA; 60 BP.
XX
AC   ABN45339;
XX
DT   15-JUL-2002  (first entry)
XX
DE   Human spliced transcript detection oligonucleotide SEQ ID NO:18087.
XX
KW   Human; mouse; rat; splice transcript; detection; RNA transcript;
KW   splice variant; transcriptome; oligonucleotide library; ss.
XX
OS   Homo sapiens.
XX
PN   WO200210449-A2.
XX
PD   07-FEB-2002.
XX
PF   20-JUL-2001; 2001WO-IB01903.
XX
PR   28-JUL-2000; 2000US-221607P.
PR   02-MAY-2001; 2001US-287724P.
XX
PA   (COMP-) COMPUGEN INC.
XX
PI   Shoshan A,  Wasserman A,  Mintz E,  Mintz L,  Faigler S;
XX
DR   WPI; 2002-257383/30.
XX
PT   New oligonucleotide libraries comprising oligonucleotides which
PT   selectively hybridize to mRNAs transcribed from a transcription unit of
PT   a genome, useful for detecting tissue-, pathology-, and
PT   developmental-specific genes  -
XX
PS   Example 1; SEQ ID 18087; 47pp; English.
XX
CC   The present invention describes oligonucleotide libraries for detecting
CC   messenger RNAs that populate a (sub-)transcriptome, where the
CC   (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC   transcription units that populate a genome. The library comprises
CC   several oligonucleotides, each capable of hybridising selectively to a
CC   set of messenger RNAs transcribed from a given transcription unit of
CC   the genome, which encodes one or more messenger RNA splice variants.
CC   The oligonucleotide libraries are useful for detecting mRNAs from a
CC   biological sample, in expression profiling studies, in qualitatively or
CC   quantitatively characterising the corresponding transcriptome, and in
CC   detecting RNA transcripts and splice variants of human or animal
CC   transcriptomes. The libraries may also be used as specialised mini
CC   libraries to detect transcripts of a sub-transcriptome under a
CC   particular biological or pathological state, and so allowing the
CC   detection of tissue- and pathology-specific genes such as those genes
CC   only expressed in specific tissue under a specific pathological
CC   condition; to detect developmental specific genes; and to detect RNA
CC   transcripts and splice variants of a transcriptome of a patient suffering
CC   from a particular disorder. ABN27253 to ABN59589 represent
CC   oligonucleotide sequences from rats, humans and mice, which are used in
CC   the exemplification of the present invention.
CC   N.B. The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 60 BP; 14 A; 17 C; 14 G; 15 T; 0 other;

  Query Match             76.0%;  Score 15.2;  DB 24;  Length 60;
  Best Local Similarity   85.0%;  Pred. No. 1e+03;
  Matches   17;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 TCCTCTGGCTTCTGGTAGCG 20
           |||||||||| |||| ||| |
Db      34 TCCTCTGGCATCTGTTAGGG 53


RESULT 7
ABN39404
ID   ABN39404 standard; DNA; 60 BP.
XX
AC   ABN39404;
XX
DT   15-JUL-2002  (first entry)
XX
DE   Human spliced transcript detection oligonucleotide SEQ ID NO:12152.
XX
KW   Human; mouse; rat; splice transcript; detection; RNA transcript;
KW   splice variant; transcriptome; oligonucleotide library; ss.
XX
OS   Homo sapiens.
XX
PN   WO200210449-A2.
XX
PD   07-FEB-2002.
XX
PF   20-JUL-2001; 2001WO-IB01903.
XX
PR   28-JUL-2000; 2000US-221607P.
PR   02-MAY-2001; 2001US-287724P.
XX
PA   (COMP-) COMPUGEN INC.
XX
PI   Shoshan A,  Wasserman A,  Mintz E,  Mintz L,  Faigler S;
XX
DR   WPI; 2002-257383/30.
XX
PT   New oligonucleotide libraries comprising oligonucleotides which
PT   selectively hybridize to mRNAs transcribed from a transcription unit of
PT   a genome, useful for detecting tissue-, pathology-, and
PT   developmental-specific genes  -
XX
PS   Example 1; SEQ ID 12152; 47pp; English.
XX
CC   The present invention describes oligonucleotide libraries for detecting
CC   messenger RNAs that populate a (sub-)transcriptome, where the
CC   (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC   transcription units that populate a genome. The library comprises
CC   several oligonucleotides, each capable of hybridising selectively to a
CC   set of messenger RNAs transcribed from a given transcription unit of
CC   the genome, which encodes one or more messenger RNA splice variants.
CC   The oligonucleotide libraries are useful for detecting mRNAs from a
CC   biological sample, in expression profiling studies, in qualitatively or
CC   quantitatively characterising the corresponding transcriptome, and in
CC   detecting RNA transcripts and splice variants of human or animal
CC   transcriptomes. The libraries may also be used as specialised mini
CC   libraries to detect transcripts of a sub-transcriptome under a
CC   particular biological or pathological state, and so allowing the
CC   detection of tissue- and pathology-specific genes such as those genes
CC   only expressed in specific tissue under a specific pathological
```

```
CC   condition; to detect developmental specific genes; and to detect RNA
CC   transcripts and splice variants of a transcriptome of a patient suffering
CC   from a particular disorder. ABN27253 to ABN59589 represent
CC   oligonucleotide sequences from rats, humans and mice, which are used in
CC   the exemplification of the present invention.
CC   N.B. The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 60 BP; 13 A; 10 C; 21 G; 16 T; 0 other;

  Query Match             74.0%;  Score 14.8;  DB 24;  Length 60;
  Best Local Similarity   88.9%;  Pred. No. 1.6e+03;
  Matches   16;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       3 CTCTGGCTTCTGGTAGCG 20
           ||||||||| ||| ||||
Db      36 CTCTGGCTTTTGGAAGCG 53


RESULT 8
AAV51908
ID   AAV51908 standard; DNA; 23 BP.
XX
AC   AAV51908;
XX
DT   02-FEB-1999  (first entry)
XX
DE   Zea mays genome reverse PCR primer #204.
XX
KW   Polymorphic marker; allele-specific; probe; amplification; PCR primer;
KW   hybridisation; plant; hybrid certification; genetic contribution;
KW   progeny; back-cross; hybrid; ancestry; corn; ss.
XX
OS   Synthetic.
OS   Zea mays.
XX
PN   WO9824796-A1.
XX
PD   11-JUN-1998.
XX
PF   01-DEC-1997;   97WO-US21782.
XX
PR   07-MAR-1997;   97US-0813507.
PR   02-DEC-1996;   96US-0032069.
XX
PA   (AFFY-) AFFYMETRIX INC.
XX
PI   Landry BS,  Lemieux B,  Murigneux A,  Sapolsky RJ;
XX
DR   WPI; 1998-333252/29.
XX
PT   Brassica species allele-specific oligonucleotide probes and primers
PT   - useful for plant breeding
XX
PS   Example 1; Page Page 53; 65pp; English.
XX
CC   AAV51705-V52008 are reverse PCR primers used to amplify fragments of the
CC   Zea mays genome in order to detect polymorphic markers. Such markers can
CC   be used in the construction of allele-specific primers and probes for
CC   amplification or hybridisation, e.g. to determine common or disparate
CC   ancestry between 2 or more plants, to monitor the genetic contribution
CC   of an ancestral plant, to trace the progeny of proprietary plants, in
CC   certification of a hybrid plant or to identify the progeny of a
CC   back-crossed plant with an ancestral plant.
XX
SQ   Sequence 23 BP; 2 A; 7 C; 6 G; 8 T; 0 other;

  Query Match             71.0%;  Score 14.2;  DB 19;  Length 23;
  Best Local Similarity   84.2%;  Pred. No. 2.8e+03;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 TCCTCTGGCTTCTGGTAGC 19
           ||| | ||||||| |||||
Db       3 TCCACGGGCTTCTTGTAGC 21


RESULT 9
AAV50992/c
ID   AAV50992 standard; DNA; 41 BP.
XX
AC   AAV50992;
XX
DT   04-JAN-1999  (first entry)
XX
DE   Maize polymorphic marker S43G1/G4-1 DNA.
XX
KW   Polymorphic marker; allele-specific; primer; probe; amplification;
KW   hybridisation; plant; hybrid certification; genetic contribution;
KW   progeny; back-cross; hybrid; ancestry; maize; ss.
XX
OS   Zea mays.
XX
FH   Key             Location/Qualifiers
FT   variation       21
FT                   /*tag= a
FT                   /replace= "a"
FT                   /note= "polymorphism"
XX
PN   WO9824796-A1.
XX
PD   11-JUN-1998.
XX
PF   01-DEC-1997;   97WO-US21782.
XX
PR   07-MAR-1997;   97US-0813507.
PR   02-DEC-1996;   96US-0032069.
XX
PA   (AFFY-) AFFYMETRIX INC.
XX
PI   Landry BS,  Lemieux B,  Murigneux A,  Sapolsky RJ;
XX
DR   WPI; 1998-333252/29.
XX
PT   Brassica species allele-specific oligonucleotide probes and primers
PT   - useful for plant breeding
XX
PS   Claim 1; Page 43; 65pp; English.
XX
CC   This DNA sequence is a region of a Zea mays genome which contains a
CC   polymorphic marker. This sequence can be used in the construction of
CC   allele-specific primers and probes for amplification or hybridisation,
CC   e.g. to determine common or disparate ancestry between 2 or more plants,
CC   to monitor the genetic contribution of an ancestral plant, to trace the
CC   progeny of proprietary plants, in certification of a hybrid plant or to
CC   identify the progeny of a back-crossed plant with an ancestral plant.
XX
SQ   Sequence 41 BP; 9 A; 10 C; 14 G; 8 T; 0 other;

  Query Match             71.0%;  Score 14.2;  DB 19;  Length 41;
  Best Local Similarity   84.2%;  Pred. No. 2.9e+03;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 TCCTCTGGCTTCTGGTAGC 19
           ||| | ||||||| |||||
Db      34 TCCACGGGCTTCTCGTAGC 16


RESULT 10
AAV50993/c
ID   AAV50993 standard; DNA; 41 BP.
XX
AC   AAV50993;
XX
DT   04-JAN-1999  (first entry)
XX
```

```
DE   Maize polymorphic marker S43G1/G4-1 DNA #2.
XX
KW   Polymorphic marker; allele-specific; primer; probe; amplification;
KW   hybridisation; plant; hybrid certification; genetic contribution;
KW   progeny; back-cross; hybrid; ancestry; maize; ss.
XX
OS   Zea mays.
XX
FH   Key             Location/Qualifiers
FT   variation       21
FT                   /*tag= a
FT                   /replace= "a"
FT                   /note= "polymorphism"
XX
PN   WO9824796-A1.
XX
PD   11-JUN-1998.
XX
PF   01-DEC-1997;   97WO-US21782.
XX
PR   07-MAR-1997;   97US-0813507.
PR   02-DEC-1996;   96US-0032069.
XX
PA   (AFFY-) AFFYMETRIX INC.
XX
PI   Landry BS,  Lemieux B,  Murigneux A,  Sapolsky RJ;
XX
DR   WPI; 1998-333252/29.
XX
PT   Brassica species allele-specific oligonucleotide probes and primers
PT   - useful for plant breeding
XX
PS   Claim 1; Page 43; 65pp; English.
XX
CC   This DNA sequence is a region of a Zea mays genome which contains a
CC   polymorphic marker. This sequence can be used in the construction of
CC   allele-specific primers and probes for amplification or hybridisation,
CC   e.g. to determine common or disparate ancestry between 2 or more plants,
CC   to monitor the genetic contribution of an ancestral plant, to trace the
CC   progeny of proprietary plants, in certification of a hybrid plant or to
CC   identify the progeny of a back-crossed plant with an ancestral plant.
XX
SQ   Sequence 41 BP; 9 A; 10 C; 15 G; 7 T; 0 other;

  Query Match             71.0%;  Score 14.2;  DB 19;  Length 41;
  Best Local Similarity   84.2%;  Pred. No. 2.9e+03;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 TCCTCTGGCTTCTGGTAGC 19
           ||| | ||||||| |||||
Db      34 TCCACGGGCTTCTCGTAGC 16


RESULT 11
AAV50969/c
ID   AAV50969 standard; DNA; 41 BP.
XX
AC   AAV50969;
XX
DT   04-JAN-1999  (first entry)
XX
DE   Maize polymorphic marker S43G2/G6-2 DNA.
XX
KW   Polymorphic marker; allele-specific; primer; probe; amplification;
KW   hybridisation; plant; hybrid certification; genetic contribution;
KW   progeny; back-cross; hybrid; ancestry; maize; ss.
XX
OS   Zea mays.
XX
FH   Key             Location/Qualifiers
FT   variation       21
FT                   /*tag= a
FT                   /replace= "g"
FT                   /note= "polymorphism"
XX
PN   WO9824796-A1.
XX
PD   11-JUN-1998.
XX
PF   01-DEC-1997;   97WO-US21782.
XX
PR   07-MAR-1997;   97US-0813507.
PR   02-DEC-1996;   96US-0032069.
XX
PA   (AFFY-) AFFYMETRIX INC.
XX
PI   Landry BS,  Lemieux B,  Murigneux A,  Sapolsky RJ;
XX
DR   WPI; 1998-333252/29.
XX
PT   Brassica species allele-specific oligonucleotide probes and primers
PT   - useful for plant breeding
XX
PS   Claim 1; Page 43; 65pp; English.
XX
CC   This DNA sequence is a region of a Zea mays genome which contains a
CC   polymorphic marker. This sequence can be used in the construction of
CC   allele-specific primers and probes for amplification or hybridisation,
CC   e.g. to determine common or disparate ancestry between 2 or more plants,
CC   to monitor the genetic contribution of an ancestral plant, to trace the
CC   progeny of proprietary plants, in certification of a hybrid plant or to
CC   identify the progeny of a back-crossed plant with an ancestral plant.
XX
SQ   Sequence 41 BP; 10 A; 10 C; 14 G; 7 T; 0 other;

  Query Match             71.0%;  Score 14.2;  DB 19;  Length 41;
  Best Local Similarity   84.2%;  Pred. No. 2.9e+03;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 TCCTCTGGCTTCTGGTAGC 19
           ||| | ||||||| |||||
Db      34 TCCACGGGCTTCTTGTAGC 16


RESULT 12
AAV50971/c
ID   AAV50971 standard; DNA; 41 BP.
XX
AC   AAV50971;
XX
DT   04-JAN-1999  (first entry)
XX
DE   Maize polymorphic marker S43G2/G6-2B DNA.
XX
KW   Polymorphic marker; allele-specific; primer; probe; amplification;
KW   hybridisation; plant; hybrid certification; genetic contribution;
KW   progeny; back-cross; hybrid; ancestry; maize; ss.
XX
OS   Zea mays.
XX
FH   Key             Location/Qualifiers
FT   variation       21
FT                   /*tag= a
FT                   /replace= "g"
FT                   /note= "polymorphism"
XX
PN   WO9824796-A1.
XX
PD   11-JUN-1998.
XX
PF   01-DEC-1997;   97WO-US21782.
XX
PR   07-MAR-1997;   97US-0813507.
PR   02-DEC-1996;   96US-0032069.
XX
PA   (AFFY-) AFFYMETRIX INC.
```

```
XX
PI   Landry BS,  Lemieux B,  Murigneux A,  Sapolsky RJ;
XX
DR   WPI; 1998-333252/29.
XX
PT   Brassica species allele-specific oligonucleotide probes and primers
PT   - useful for plant breeding
XX
PS   Claim 1; Page 43; 65pp; English.
XX
CC   This DNA sequence is a region of a Zea mays genome which contains a
CC   polymorphic marker. This sequence can be used in the construction of
CC   allele-specific primers and probes for amplification or hybridisation,
CC   e.g. to determine common or disparate ancestry between 2 or more plants,
CC   to monitor the genetic contribution of an ancestral plant, to trace the
CC   progeny of proprietary plants, in certification of a hybrid plant or to
CC   identify the progeny of a back-crossed plant with an ancestral plant.
XX
SQ   Sequence 41 BP; 10 A; 10 C; 13 G; 8 T; 0 other;

  Query Match             71.0%;  Score 14.2;  DB 19;  Length 41;
  Best Local Similarity   84.2%;  Pred. No. 2.9e+03;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 TCCTCTGGCTTCTGGTAGC 19
           ||| | ||||||| |||||
Db      34 TCCACGGGCTTCTTGTAGC 16


RESULT 13
AAV50980/c
ID   AAV50980 standard; DNA; 41 BP.
XX
AC   AAV50980;
XX
DT   04-JAN-1999  (first entry)
XX
DE   Maize polymorphic marker S43G2/G6-2 DNA.
XX
KW   Polymorphic marker; allele-specific; primer; probe; amplification;
KW   hybridisation; plant; hybrid certification; genetic contribution;
KW   progeny; back-cross; hybrid; ancestry; maize; ss.
XX
OS   Zea mays.
XX
FH   Key             Location/Qualifiers
FT   variation       21
FT                   /*tag= a
FT                   /replace= "g"
FT                   /note= "polymorphism"
XX
PN   WO9824796-A1.
XX
PD   11-JUN-1998.
XX
PF   01-DEC-1997;   97WO-US21782.
XX
PR   07-MAR-1997;   97US-0813507.
PR   02-DEC-1996;   96US-0032069.
XX
PA   (AFFY-) AFFYMETRIX INC.
XX
PI   Landry BS,  Lemieux B,  Murigneux A,  Sapolsky RJ;
XX
DR   WPI; 1998-333252/29.
XX
PT   Brassica species allele-specific oligonucleotide probes and primers
PT   - useful for plant breeding
XX
PS   Claim 1; Page 43; 65pp; English.
XX
CC   This DNA sequence is a region of a Zea mays genome which contains a
CC   polymorphic marker. This sequence can be used in the construction of
CC   allele-specific primers and probes for amplification or hybridisation,
CC   e.g. to determine common or disparate ancestry between 2 or more plants,
CC   to monitor the genetic contribution of an ancestral plant, to trace the
CC   progeny of proprietary plants, in certification of a hybrid plant or to
CC   identify the progeny of a back-crossed plant with an ancestral plant.
XX
SQ   Sequence 41 BP; 10 A; 10 C; 14 G; 7 T; 0 other;

  Query Match             71.0%;  Score 14.2;  DB 19;  Length 41;
  Best Local Similarity   84.2%;  Pred. No. 2.9e+03;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 TCCTCTGGCTTCTGGTAGC 19
           ||| | ||||||| |||||
Db      34 TCCACGGGCTTCTTGTAGC 16


RESULT 14
AAV50981/c
ID   AAV50981 standard; DNA; 41 BP.
XX
AC   AAV50981;
XX
DT   04-JAN-1999  (first entry)
XX
DE   Maize polymorphic marker S43G2/G6-2B DNA.
XX
KW   Polymorphic marker; allele-specific; primer; probe; amplification;
KW   hybridisation; plant; hybrid certification; genetic contribution;
KW   progeny; back-cross; hybrid; ancestry; maize; ss.
XX
OS   Zea mays.
XX
FH   Key             Location/Qualifiers
FT   variation       21
FT                   /*tag= a
FT                   /replace= "g"
FT                   /note= "polymorphism"
XX
PN   WO9824796-A1.
XX
PD   11-JUN-1998.
XX
PF   01-DEC-1997;   97WO-US21782.
XX
PR   07-MAR-1997;   97US-0813507.
PR   02-DEC-1996;   96US-0032069.
XX
PA   (AFFY-) AFFYMETRIX INC.
XX
PI   Landry BS,  Lemieux B,  Murigneux A,  Sapolsky RJ;
XX
DR   WPI; 1998-333252/29.
XX
PT   Brassica species allele-specific oligonucleotide probes and primers
PT   - useful for plant breeding
XX
PS   Claim 1; Page 43; 65pp; English.
XX
CC   This DNA sequence is a region of a Zea mays genome which contains a
CC   polymorphic marker. This sequence can be used in the construction of
CC   allele-specific primers and probes for amplification or hybridisation,
CC   e.g. to determine common or disparate ancestry between 2 or more plants,
CC   to monitor the genetic contribution of an ancestral plant, to trace the
CC   progeny of proprietary plants, in certification of a hybrid plant or to
CC   identify the progeny of a back-crossed plant with an ancestral plant.
XX
SQ   Sequence 41 BP; 10 A; 10 C; 13 G; 8 T; 0 other;

  Query Match             71.0%;  Score 14.2;  DB 19;  Length 41;
  Best Local Similarity   84.2%;  Pred. No. 2.9e+03;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;
```

```
Qy      1 TCCTCTGGCTTCTGGTAGC 19
          ||| | ||||||| |||||
Db     34 TCCACGGGCTTCTTGTAGC 16
```

RESULT 15
AAV47809/c
ID AAV47809 standard; DNA; 41 BP.
XX
AC AAV47809;
XX
DT 14-OCT-1998 (first entry)
XX
DE Maize polymorphic site oligonucleotide marker Wx1-G1/G4-1.
XX
KW Maize; marker; probe; PCR primer; polymorphism; vegetal sequence;
KW polymorphic site; corn; graminae species; ss.
XX
OS Synthetic.
OS Zea sp.
XX
PN WO9830717-A2.
XX
PD 16-JUL-1998.
XX
PF 02-DEC-1997; 97WO-EP07134.
XX
PR 02-DEC-1996; 96US-0032069.
XX
PA (BIOC-) BIOCEM SA.
XX
PI Murigneux A;
XX
DR WPI; 1998-399160/34.
XX
PT Vegetal sequences including single nucleotide polymorphism - useful,
PT e.g. to determine polymorphisms in plants, determine strain in plant
PT breeding and to correlate polymorphisms with phenotypic traits
XX
PS Claim 2; Page 13; 32pp; English.
XX
CC The present invention describes a nucleic acid segment comprising at
CC least 10 contiguous nucleotides from a vegetal sequence including a
CC polymorphic site which is a single nucleotide polymorphism (SNP), or the
CC complement of the segment. Also described are: (1) an allele-specific
CC oligonucleotides hybridising to segment, or their complements, and (2) a
CC method of analysing nucleic acids from a subject, by determining if a
CC base is occupying any one (or a set) of polymorphic sites in 261
CC sequences derived from six maize lines (see AAV47701 to AAV47961). The
CC segments are useful in fingerprint analysis in plants to determine which
CC polymorphisms are present, which strain a plant belongs to and to
CC distinguish between strains. The polymorphisms may correlate with
CC phenotypic traits (e.g. plant growth rate or crop yield), and the
CC segments are useful to determine the presence/absence of specific
CC polymorphisms correlating with the existence/absence of particular
CC traits. The segments are also useful in marker assisted back-cross
CC techniques to select plants with a higher percentage of recurrent parent
CC in a back-cross population. Segments incorporate SNPs which occur more
CC frequently than other polymorphism types and are therefore more likely
CC to be located close to genetic loci of interest; different forms of
CC characterised SNPs are also often easier to detect than other
CC polymorphism types.
XX
SQ Sequence 41 BP; 9 A; 10 C; 13 G; 8 T; 1 other;

```
  Query Match             71.0%;  Score 14.2;  DB 19;  Length 41;
  Best Local Similarity   84.2%;  Pred. No. 2.9e+03;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 TCCTCTGGCTTCTGGTAGC 19
          ||| | ||||||| |||||
Db     34 TCCACGGGCTTCTYGTAGC 16
```

Search completed: November 23, 2002, 06:29:18
Job time : 101.6 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 05:53:51 ; Search time 21.55 Seconds
(without alignments)
284.619 Million cell updates/sec

Title: US-09-296-264-16
Perfect score: 20
Sequence: 1 tcctctggcttctggtagcg 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 687286

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
c	1	15.4	77.0	50	1	US-08-171-389-196	Sequence 196, App
c	2	15.4	77.0	50	1	US-08-123-936-196	Sequence 196, App
c	3	15.4	77.0	50	2	US-08-475-228A-196	Sequence 196, App
c	4	15.4	77.0	50	3	US-08-482-080A-196	Sequence 196, App
c	5	15.4	77.0	50	4	US-09-354-947-196	Sequence 196, App
c	6	15.4	77.0	50	5	PCT-US93-12388-196	Sequence 196, App
	7	13.6	68.0	51	4	US-09-548-372D-41	Sequence 41, Appl
c	8	13.6	68.0	51	4	US-09-548-372D-42	Sequence 42, Appl
	9	13.6	68.0	51	4	US-09-548-367D-41	Sequence 41, Appl
c	10	13.6	68.0	51	4	US-09-548-367D-42	Sequence 42, Appl
c	11	13.4	67.0	54	1	US-08-758-306-1360	Sequence 1360, Ap
c	12	13.2	66.0	23	2	US-08-742-297-5	Sequence 5, Appli
	13	13.2	66.0	73	2	US-07-916-098A-47	Sequence 47, Appl
c	14	13.2	66.0	77	2	US-07-916-098A-46	Sequence 46, Appl
c	15	13.2	66.0	77	2	US-07-916-098A-48	Sequence 48, Appl
c	16	12.8	64.0	22	3	US-08-784-582-64	Sequence 64, Appl
	17	12.8	64.0	60	1	US-08-478-039-9	Sequence 9, Appli
	18	12.8	64.0	60	1	US-08-476-349A-9	Sequence 9, Appli
c	19	12.6	63.0	32	3	US-08-777-266A-1	Sequence 1, Appli
c	20	12.6	63.0	32	4	US-09-326-186B-1	Sequence 1, Appli
c	21	12.6	63.0	35	1	US-08-110-300A-17	Sequence 17, Appl
c	22	12.6	63.0	35	2	US-08-886-642-17	Sequence 17, Appl
	23	12.6	63.0	39	3	US-08-997-918-35	Sequence 35, Appl
	24	12.6	63.0	56	2	US-07-673-661B-5	Sequence 5, Appli
c	25	12.6	63.0	68	1	US-08-208-886C-23	Sequence 23, Appl
c	26	12.6	63.0	68	1	US-08-704-744-23	Sequence 23, Appl
c	27	12.6	63.0	68	1	US-08-469-557-23	Sequence 23, Appl
c	28	12.6	63.0	68	2	US-08-290-793B-23	Sequence 23, Appl
c	29	12.6	63.0	84	4	US-09-250-609-16	Sequence 16, Appl
c	30	12.4	62.0	17	4	US-08-584-040-4182	Sequence 4182, Ap
c	31	12.4	62.0	21	3	US-09-215-966-18	Sequence 18, Appl
c	32	12.4	62.0	46	4	US-09-052-521C-29	Sequence 29, Appl
c	33	12.4	62.0	61	5	PCT-US91-02942-15	Sequence 15, Appl
	34	12.4	62.0	100	4	US-09-298-886-23	Sequence 23, Appl
	35	12.4	62.0	100	4	US-09-298-886-26	Sequence 26, Appl
c	36	12.2	61.0	19	3	US-08-746-411A-1	Sequence 1, Appli
c	37	12.2	61.0	19	4	US-08-857-046A-1	Sequence 1, Appli
	38	12.2	61.0	20	4	US-09-011-745-16	Sequence 16, Appl
c	39	12.2	61.0	23	1	US-08-056-200-45	Sequence 45, Appl
c	40	12.2	61.0	23	2	US-08-800-644-45	Sequence 45, Appl
c	41	12.2	61.0	28	2	US-08-811-492-60	Sequence 60, Appl
c	42	12.2	61.0	28	5	PCT-US96-10545A-60	Sequence 60, Appl
	43	12.2	61.0	62	4	US-09-453-702B-114	Sequence 114, App
c	44	12.2	61.0	71	1	US-08-472-255A-82	Sequence 82, Appl
c	45	12.2	61.0	71	1	US-08-479-724A-82	Sequence 82, Appl

ALIGNMENTS

```
RESULT 1
US-08-171-389-196/c
; Sequence 196, Application US/08171389
; Patent No. 5578444
;  GENERAL INFORMATION:
;    APPLICANT:  Edwards, Cynthia A.
;    APPLICANT:  Cantor, Charles R.
;    APPLICANT:  Andrews, Beth M.
;    APPLICANT:  Turin, Lisa M.
;    APPLICANT:  Fry, Kirk E.
;    TITLE OF INVENTION:  Sequence-Directed DNA Binding
;    TITLE OF INVENTION:  Molecules, Compositions and Methods
;    NUMBER OF SEQUENCES:  641
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genelabs Technologies, Inc.
;      STREET:  505 Penobscot Drive
;      CITY:  Redwood City
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94063
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/171,389
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/123,936
;      FILING DATE:  17-SEP-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/996,783
;      FILING DATE:  23-DEC-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/723,618
;      FILING DATE:  27-JUN-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/081,070
;      FILING DATE:  22-JUN-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fabian, Gary R.
;      REGISTRATION NUMBER:  33,875
;      REFERENCE/DOCKET NUMBER:  4600-0175/G19P3
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 324-0880
;      TELEFAX:  (415) 324-0960
;  INFORMATION FOR SEQ ID NO:  196:
;    SEQUENCE CHARACTERISTICS:
```

```
;      LENGTH:  50 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ORIGINAL SOURCE:
;      INDIVIDUAL ISOLATE:  Human immune interferon (INF-gamma)
;      INDIVIDUAL ISOLATE:  gene
US-08-171-389-196

  Query Match             77.0%;  Score 15.4;  DB 1;  Length 50;
  Best Local Similarity   94.1%;  Pred. No. 1.1e+02;
  Matches   16;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 TCCTCTGGCTTCTGGTA 17
          ||||||||||| ||||||
Db     45 TCCTCTGGCTGCTGGTA 29


RESULT 2
US-08-123-936-196/c
; Sequence 196, Application US/08123936
; Patent No. 5726014
;  GENERAL INFORMATION:
;    APPLICANT:  Edwards, Cynthia A.
;    APPLICANT:  Cantor, Charles R.
;    APPLICANT:  Andrews, Beth M.
;    APPLICANT:  Turin, Lisa M.
;    TITLE OF INVENTION:  Screening Assay for the Detection of
;    TITLE OF INVENTION:  DNA-Binding Molecules
;    NUMBER OF SEQUENCES:  640
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genelabs Technologies, Inc.
;      STREET:  505 Penobscot Drive
;      CITY:  Redwood City
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94063
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/123,936
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/996,783
;      FILING DATE:  23-DEC-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/723,618
;      FILING DATE:  27-JUN-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fabian, Gary R.
;      REGISTRATION NUMBER:  33,875
;      REFERENCE/DOCKET NUMBER:  4600-0075.32/G19P2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 324-0880
;      TELEFAX:  (415) 324-0960
;  INFORMATION FOR SEQ ID NO:  196:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  50 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ORIGINAL SOURCE:
;      INDIVIDUAL ISOLATE:  Human immune interferon (INF-gamma)
;      INDIVIDUAL ISOLATE:  gene
US-08-123-936-196

  Query Match             77.0%;  Score 15.4;  DB 1;  Length 50;
  Best Local Similarity   94.1%;  Pred. No. 1.1e+02;
  Matches   16;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 TCCTCTGGCTTCTGGTA 17
          ||||||||||| ||||||
Db     45 TCCTCTGGCTGCTGGTA 29


RESULT 3
US-08-475-228A-196/c
; Sequence 196, Application US/08475228A
; Patent No. 5869241
;  GENERAL INFORMATION:
;    APPLICANT:  Edwards, Cynthia A.
;    APPLICANT:  Cantor, Charles R.
;    APPLICANT:  Andrews, Beth M.
;    APPLICANT:  Turin, Lisa M.
;    APPLICANT:  Fry, Kirk E.
;    TITLE OF INVENTION:  Sequence-Directed DNA Binding
;    TITLE OF INVENTION:  Molecules, Compositions and Methods
;    NUMBER OF SEQUENCES:  664
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genelabs Technologies, Inc.
;      STREET:  505 Penobscot Drive
;      CITY:  Redwood City
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94063
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/475,228A
;      FILING DATE:  06-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/123,936
;      FILING DATE:  17-SEP-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/996,783
;      FILING DATE:  23-DEC-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/723,618
;      FILING DATE:  27-JUN-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/081,070
;      FILING DATE:  22-JUN-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Stratford, Carol A.
;      REGISTRATION NUMBER:  34,444
;      REFERENCE/DOCKET NUMBER:  4600-0175.21/G19P3D2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 324-0880
;      TELEFAX:  (415) 324-0960
;  INFORMATION FOR SEQ ID NO:  196:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  50 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ORIGINAL SOURCE:
;      INDIVIDUAL ISOLATE:  Human immune interferon (INF-gamma)
;      INDIVIDUAL ISOLATE:  gene
US-08-475-228A-196

  Query Match             77.0%;  Score 15.4;  DB 2;  Length 50;
  Best Local Similarity   94.1%;  Pred. No. 1.1e+02;
  Matches   16;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;
```

```
Qy      1 TCCTCTGGCTTCTGGTA 17
          ||||||||||  ||||||
Db     45 TCCTCTGGCTGCTGGTA 29


RESULT 4
US-08-482-080A-196/c
; Sequence 196, Application US/08482080A
; Patent No. 6010849
;  GENERAL INFORMATION:
;    APPLICANT:  Edwards, Cynthia A.
;    APPLICANT:  Cantor, Charles R.
;    APPLICANT:  Andrews, Beth M.
;    APPLICANT:  Turin, Lisa M.
;    APPLICANT:  Fry, Kirk E.
;    TITLE OF INVENTION:  Sequence-Directed DNA Binding
;    TITLE OF INVENTION:  Molecules, Compositions and Methods
;    NUMBER OF SEQUENCES:  664
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genelabs Technologies, Inc.
;      STREET:  505 Penobscot Drive
;      CITY:  Redwood City
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94063
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/482,080A
;      FILING DATE:  07-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/171,389
;      FILING DATE:  20-DEC-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/123,936
;      FILING DATE:  17-SEP-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/996,783
;      FILING DATE:  23-DEC-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/723,618
;      FILING DATE:  27-JUN-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/081,070
;      FILING DATE:  22-JUN-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Brady, John F.
;      REGISTRATION NUMBER:  39,118
;      REFERENCE/DOCKET NUMBER:  4600-0175.20/G19P3D1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (650) 324-0880
;      TELEFAX:  (650) 324-0960
;  INFORMATION FOR SEQ ID NO:  196:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  50 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ORIGINAL SOURCE:
;      INDIVIDUAL ISOLATE:  Human immune interferon (INF-gamma)
;      INDIVIDUAL ISOLATE:  gene
US-08-482-080A-196

  Query Match            77.0%;  Score 15.4;  DB 3;  Length 50;
  Best Local Similarity  94.1%;  Pred. No. 1.1e+02;
  Matches   16;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;
```

```
Qy      1 TCCTCTGGCTTCTGGTA 17
          ||||||||||  ||||||
Db     45 TCCTCTGGCTGCTGGTA 29


RESULT 5
US-09-354-947-196/c
; Sequence 196, Application US/09354947
; Patent No. 6384208
;  GENERAL INFORMATION:
;    APPLICANT:  Edwards, Cynthia A.
;    APPLICANT:  Cantor, Charles R.
;    APPLICANT:  Andrews, Beth M.
;    APPLICANT:  Turin, Lisa M.
;    APPLICANT:  Fry, Kirk E.
;    TITLE OF INVENTION:  Sequence-Directed DNA Binding
;    TITLE OF INVENTION:  Molecules, Compositions and Methods
;    NUMBER OF SEQUENCES:  664
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genelabs Technologies, Inc.
;      STREET:  505 Penobscot Drive
;      CITY:  Redwood City
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94063
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/354,947
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/482,080
;      FILING DATE:  07-JUN-1995
;      APPLICATION NUMBER:  US 08/171,389
;      FILING DATE:  20-DEC-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/123,936
;      FILING DATE:  17-SEP-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/996,783
;      FILING DATE:  23-DEC-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/723,618
;      FILING DATE:  27-JUN-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/081,070
;      FILING DATE:  22-JUN-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Brady, John F.
;      REGISTRATION NUMBER:  39,118
;      REFERENCE/DOCKET NUMBER:  4600-0175.20/G19P3D1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (650) 324-0880
;      TELEFAX:  (650) 324-0960
;  INFORMATION FOR SEQ ID NO:  196:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  50 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ORIGINAL SOURCE:
;      INDIVIDUAL ISOLATE:  Human immune interferon (INF-gamma)
;      INDIVIDUAL ISOLATE:  gene
US-09-354-947-196

  Query Match            77.0%;  Score 15.4;  DB 4;  Length 50;
  Best Local Similarity  94.1%;  Pred. No. 1.1e+02;
  Matches   16;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;
```

```
Qy       1 TCCTCTGGCTTCTGGTA 17
           ||||||||| ||||||
Db      45 TCCTCTGGCTGCTGGTA 29


RESULT 6
PCT-US93-12388-196/c
; Sequence 196, Application PC/TUS9312388
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  Sequence-Directed DNA Binding
;    TITLE OF INVENTION:  Molecules, Compositions and Methods
;    NUMBER OF SEQUENCES:  641
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genelabs Technologies, Inc.
;      STREET:  505 Penobscot Drive
;      CITY:  Redwood City
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94063
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US93/12388
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/123,936
;      FILING DATE:  17-SEP-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/996,783
;      FILING DATE:  23-DEC-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fabian, Gary R.
;      REGISTRATION NUMBER:  33,875
;      REFERENCE/DOCKET NUMBER:  4600-0175.41/G19PCT2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 324-0880
;      TELEFAX:  (415) 324-0960
;  INFORMATION FOR SEQ ID NO:  196:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  50 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ORIGINAL SOURCE:
;      INDIVIDUAL ISOLATE:  Human immune interferon (INF-gamma)
;      INDIVIDUAL ISOLATE:  gene
PCT-US93-12388-196

  Query Match             77.0%;  Score 15.4;  DB 5;  Length 50;
  Best Local Similarity   94.1%;  Pred. No. 1.1e+02;
  Matches   16;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 TCCTCTGGCTTCTGGTA 17
           |||||||||| ||||||
Db      45 TCCTCTGGCTGCTGGTA 29


RESULT 7
US-09-548-372D-41
; Sequence 41, Application US/09548372D
; Patent No. 6420534
; GENERAL INFORMATION:
;  APPLICANT: GURNEY ET AL.
;  TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USES
;  TITLE OF INVENTION:  THEREOF
;  FILE REFERENCE: 29915/6280I
;  CURRENT APPLICATION NUMBER: US/09/548,372D
;  CURRENT FILING DATE:  2000-04-12
;  PRIOR APPLICATION NUMBER: US 60/155,493
;  PRIOR FILING DATE: 1999-09-23
;  PRIOR APPLICATION NUMBER: US 09/404,133
;  PRIOR FILING DATE: 1999-09-23
;  PRIOR APPLICATION NUMBER: PCT/US99/20881
;  PRIOR FILING DATE: 1999-09-23
;  PRIOR APPLICATION NUMBER: US 60/101,594
;  PRIOR FILING DATE: 1998-09-24
;  NUMBER OF SEQ ID NOS: 73
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 41
;   LENGTH: 51
;   TYPE: DNA
;   ORGANISM: Caspase-8 Cleavage Site
US-09-548-372D-41

  Query Match             68.0%;  Score 13.6;  DB 4;  Length 51;
  Best Local Similarity   80.0%;  Pred. No. 7.3e+02;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 TCCTCTGGCTTCTGGTAGCG 20
           ||| ||||  ||||||| ||
Db      23 TCCGCTGGACTCTGGTATCG 42


RESULT 8
US-09-548-372D-42/c
; Sequence 42, Application US/09548372D
; Patent No. 6420534
; GENERAL INFORMATION:
;  APPLICANT: GURNEY ET AL.
;  TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USE
;  TITLE OF INVENTION:  THEREOF
;  FILE REFERENCE: 29915/6280I
;  CURRENT APPLICATION NUMBER: US/09/548,372D
;  CURRENT FILING DATE:  2000-04-12
;  PRIOR APPLICATION NUMBER: US 60/155,493
;  PRIOR FILING DATE: 1999-09-23
;  PRIOR APPLICATION NUMBER: US 09/404,133
;  PRIOR FILING DATE: 1999-09-23
;  PRIOR APPLICATION NUMBER: PCT/US99/20881
;  PRIOR FILING DATE: 1999-09-23
;  PRIOR APPLICATION NUMBER: US 60/101,594
;  PRIOR FILING DATE: 1998-09-24
;  NUMBER OF SEQ ID NOS: 73
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 42
;   LENGTH: 51
;   TYPE: DNA
;   ORGANISM: Caspase-8 Cleavage Site
US-09-548-372D-42

  Query Match             68.0%;  Score 13.6;  DB 4;  Length 51;
  Best Local Similarity   80.0%;  Pred. No. 7.3e+02;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 TCCTCTGGCTTCTGGTAGCG 20
           ||| ||||  ||||||| ||
Db      33 TCCGCTGGACTCTGGTATCG 14


RESULT 9
US-09-548-367D-41
; Sequence 41, Application US/09548367D
; Patent No. 6440698
; GENERAL INFORMATION:
;  APPLICANT: GURNEY ET AL.
;  TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USE
;  TITLE OF INVENTION:  THEREOF
;  FILE REFERENCE: 29915/6280H
```

```
;  CURRENT APPLICATION NUMBER: US/09/548,367D
;  CURRENT FILING DATE:  2000-04-12
;  PRIOR APPLICATION NUMBER: US 60/155,493
;  PRIOR FILING DATE: 1999-09-23
;  PRIOR APPLICATION NUMBER: US 09/404,133
;  PRIOR FILING DATE: 1999-09-23
;  PRIOR APPLICATION NUMBER: PCT/US99/20881
;  PRIOR FILING DATE: 1999-09-23
;  PRIOR APPLICATION NUMBER: US 60/101,594
;  PRIOR FILING DATE: 1998-09-24
;  NUMBER OF SEQ ID NOS: 73
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 41
;   LENGTH: 51
;   TYPE: DNA
;   ORGANISM: Caspase-8 Cleavage Site
US-09-548-367D-41

  Query Match             68.0%;  Score 13.6;  DB 4;  Length 51;
  Best Local Similarity   80.0%;  Pred. No. 7.3e+02;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 TCCTCTGGCTTCTGGTAGCG 20
           ||| ||||  ||||||| ||
Db      23 TCCGCTGGACTCTGGTATCG 42


RESULT 10
US-09-548-367D-42/c
; Sequence 42, Application US/09548367D
; Patent No. 6440698
; GENERAL INFORMATION:
;  APPLICANT: GURNEY ET AL.
;  TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USES
;  TITLE OF INVENTION:  THEREOF
;  FILE REFERENCE: 29915/6280H
;  CURRENT APPLICATION NUMBER: US/09/548,367D
;  CURRENT FILING DATE:  2000-04-12
;  PRIOR APPLICATION NUMBER: US 60/155,493
;  PRIOR FILING DATE: 1999-09-23
;  PRIOR APPLICATION NUMBER: US 09/404,133
;  PRIOR FILING DATE: 1999-09-23
;  PRIOR APPLICATION NUMBER: PCT/US99/20881
;  PRIOR FILING DATE: 1999-09-23
;  PRIOR APPLICATION NUMBER: US 60/101,594
;  PRIOR FILING DATE: 1998-09-24
;  NUMBER OF SEQ ID NOS: 73
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 42
;   LENGTH: 51
;   TYPE: DNA
;   ORGANISM: Caspase-8 Cleavage Site
US-09-548-367D-42

  Query Match             68.0%;  Score 13.6;  DB 4;  Length 51;
  Best Local Similarity   80.0%;  Pred. No. 7.3e+02;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 TCCTCTGGCTTCTGGTAGCG 20
           ||| ||||  ||||||| ||
Db      33 TCCGCTGGACTCTGGTATCG 14


RESULT 11
US-08-758-306-1360/c
; Sequence 1360, Application US/08758306
; Patent No. 5807743
;  GENERAL INFORMATION:
;    APPLICANT:  Stinchcomb, Dan T.
;    APPLICANT:  McSwiggen, James A.
;    TITLE OF INVENTION:  METHOD AND REAGENT FOR THE
;    TITLE OF INVENTION:  TREATMENT OF DISEASES
;    TITLE OF INVENTION:  ASSOCIATED WITH
;    TITLE OF INVENTION:  INTERLEUKIN-2 RECEPTOR
;    TITLE OF INVENTION:  GAMMA-CHAIN EXPRESSION
;    NUMBER OF SEQUENCES:  1379
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Lyon & Lyon
;      STREET:  633 West Fifth Street
;      STREET:  Suite 4700
;      CITY:  Los Angeles
;      STATE:  California
;      COUNTRY:  U.S.A.
;      ZIP:  90071-2066
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" Diskette, 1.44 Mb
;      MEDIUM TYPE:  storage
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  IBM P.C. DOS 5.0
;      SOFTWARE:  FastSeq Version 1.5
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/758,306
;      FILING DATE:  December 3, 1996
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Warburg, Richard J.
;      REGISTRATION NUMBER:  32,327
;      REFERENCE/DOCKET NUMBER:  212/132
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (213) 489-1600
;      TELEFAX:  (213) 955-0440
;      TELEX:  67-3510
;  INFORMATION FOR SEQ ID NO:  1360:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  54 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-758-306-1360

  Query Match             67.0%;  Score 13.4;  DB 1;  Length 54;
  Best Local Similarity   93.3%;  Pred. No. 9.1e+02;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       5 CTGGCTTCTGGTAGC 19
           ||||||||||| ||
Db      17 CTGGCTTCTGGTTGC 3


RESULT 12
US-08-742-297-5/c
; Sequence 5, Application US/08742297
; Patent No. 5863731
;  GENERAL INFORMATION:
;    APPLICANT:  Romaine, C. Peter
;    APPLICANT:  Moorman, Gary W.
;    APPLICANT:  Sulzinski, Michael A.
;    TITLE OF INVENTION:  A Polymerase Chain Reaction-based
;    TITLE OF INVENTION:  Diagnostic Test for the Detection of Xanthomonas
;    TITLE OF INVENTION:  campestris pathovar pelargonii
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  McDonnell Boehnen Hulbert & Berghoff, Ltd.
;      STREET:  300 South Wacker Drive Seventh Floor
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60606
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
```

```
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/742,297
;      FILING DATE:  01-NOV-1996
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  No. 5863731nan, Kevin E
;      REGISTRATION NUMBER:  35,303
;      REFERENCE/DOCKET NUMBER:  96,2025
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312-913-0001
;      TELEFAX:  312-913-9808
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  23 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-742-297-5

  Query Match             66.0%;  Score 13.2;  DB 2;  Length 23;
  Best Local Similarity   83.3%;  Pred. No. 1e+03;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       3 CTCTGGCTTCTGGTAGCG 20
           || || ||| ||||||||
Db      21 CTTTGCCTTTTGGTAGCG 4


RESULT 13
US-07-916-098A-47
; Sequence 47, Application US/07916098A
; Patent No. 5871732
;  GENERAL INFORMATION:
;    APPLICANT:  BURKLY, LINDA C.
;    APPLICANT:  CHISHOLM, PATRICIA L.
;    APPLICANT:  THOMAS, DAVID W.
;    APPLICANT:  ROSA, MARGARET D.
;    APPLICANT:  ROSA, JOSEPH J.
;    TITLE OF INVENTION:  ANTI-CD4 ANTIBODY HOMOLOGS USEFUL IN
;    TITLE OF INVENTION:  PROPHYLAXIS AND TREATMENT OF AIDS, ARC AND HIV INFECTION
;    NUMBER OF SEQUENCES:  61
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  ALLEGRETTI & WITCOFF, LTD.
;      STREET:  10 SOUTH WACKER DRIVE
;      CITY:  CHICAGO
;      STATE:  ILLINOIS
;      COUNTRY:  U.S.A.
;      ZIP:  60606
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  WORD PERFECT 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/916,098A
;      FILING DATE:  July 24, 1992
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US91/08843
;      FILING DATE:  No. 5871732ember 27, 1991
;      CLASSIFICATION:  424
;      APPLICATION NUMBER:  07/618,542
;      FILING DATE:  No. 5871732ember 27, 1990
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  JOHN J. MC DONNELL
;      REGISTRATION NUMBER:  26,949
;      REFERENCE/DOCKET NUMBER:  92,310-G
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (312) 715-1000
;      TELEFAX:  (312) 715-1234
;      TELEX:  910/221-5317
;  INFORMATION FOR SEQ ID NO:  47:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  73 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  1
;      OTHER INFORMATION:  /note= "360-82 oligonucleotide"
US-07-916-098A-47

  Query Match             66.0%;  Score 13.2;  DB 2;  Length 73;
  Best Local Similarity   83.3%;  Pred. No. 1.1e+03;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 CCTCTGGCTTCTGGTAGC 19
           |||||||| |||| |||
Db       5 CCTCTGGCACCTGGCAGC 22


RESULT 14
US-07-916-098A-46/c
; Sequence 46, Application US/07916098A
; Patent No. 5871732
;  GENERAL INFORMATION:
;    APPLICANT:  BURKLY, LINDA C.
;    APPLICANT:  CHISHOLM, PATRICIA L.
;    APPLICANT:  THOMAS, DAVID W.
;    APPLICANT:  ROSA, MARGARET D.
;    APPLICANT:  ROSA, JOSEPH J.
;    TITLE OF INVENTION:  ANTI-CD4 ANTIBODY HOMOLOGS USEFUL IN
;    TITLE OF INVENTION:  PROPHYLAXIS AND TREATMENT OF AIDS, ARC AND HIV INFECTION
;    NUMBER OF SEQUENCES:  61
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  ALLEGRETTI & WITCOFF, LTD.
;      STREET:  10 SOUTH WACKER DRIVE
;      CITY:  CHICAGO
;      STATE:  ILLINOIS
;      COUNTRY:  U.S.A.
;      ZIP:  60606
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  WORD PERFECT 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/916,098A
;      FILING DATE:  July 24, 1992
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US91/08843
;      FILING DATE:  No. 5871732ember 27, 1991
;      CLASSIFICATION:  424
;      APPLICATION NUMBER:  07/618,542
;      FILING DATE:  No. 5871732ember 27, 1990
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  JOHN J. MC DONNELL
;      REGISTRATION NUMBER:  26,949
;      REFERENCE/DOCKET NUMBER:  92,310-G
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (312) 715-1000
;      TELEFAX:  (312) 715-1234
;      TELEX:  910/221-5317
;  INFORMATION FOR SEQ ID NO:  46:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  77 base pairs
;      TYPE:  nucleic acid
```

```
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  1
;      OTHER INFORMATION:  /note= "360-81 oligonucleotide"
US-07-916-098A-46

  Query Match             66.0%;  Score 13.2;  DB 2;  Length 77;
  Best Local Similarity   83.3%;  Pred. No. 1.1e+03;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 CCTCTGGCTTCTGGTAGC 19
           ||||||||  |||| |||
Db      73 CCTCTGGCACCTGGCAGC 56


RESULT 15
US-07-916-098A-48/c
; Sequence 48, Application US/07916098A
; Patent No. 5871732
;  GENERAL INFORMATION:
;    APPLICANT:  BURKLY, LINDA C.
;    APPLICANT:  CHISHOLM, PATRICIA L.
;    APPLICANT:  THOMAS, DAVID W.
;    APPLICANT:  ROSA, MARGARET D.
;    APPLICANT:  ROSA, JOSEPH J.
;    TITLE OF INVENTION:  ANTI-CD4 ANTIBODY HOMOLOGS USEFUL IN
;    TITLE OF INVENTION:  PROPHYLAXIS AND TREATMENT OF AIDS, ARC AND HIV INFECTION
;    NUMBER OF SEQUENCES:  61
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  ALLEGRETTI & WITCOFF, LTD.
;      STREET:  10 SOUTH WACKER DRIVE
;      CITY:  CHICAGO
;      STATE:  ILLINOIS
;      COUNTRY:  U.S.A.
;      ZIP:  60606
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  WORD PERFECT 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/916,098A
;      FILING DATE:  July 24, 1992
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US91/08843
;      FILING DATE:  No. 5871732ember 27, 1991
;      CLASSIFICATION:  424
;      APPLICATION NUMBER:  07/618,542
;      FILING DATE:  No. 5871732ember 27, 1990
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  JOHN J. MC DONNELL
;      REGISTRATION NUMBER:  26,949
;      REFERENCE/DOCKET NUMBER:  92,310-G
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (312) 715-1000
;      TELEFAX:  (312) 715-1234
;      TELEX:  910/221-5317
;  INFORMATION FOR SEQ ID NO:  48:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  77 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  1
;      OTHER INFORMATION:  /note= "pMDR985 insert"
US-07-916-098A-48

  Query Match             66.0%;  Score 13.2;  DB 2;  Length 77;
  Best Local Similarity   83.3%;  Pred. No. 1.1e+03;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 CCTCTGGCTTCTGGTAGC 19
           ||||||||  |||| |||
Db      73 CCTCTGGCACCTGGCAGC 56


Search completed: November 23, 2002, 06:36:20
Job time : 23.55 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 05:54:41 ; Search time 17.25 Seconds
(without alignments)
439.108 Million cell updates/sec

Title: US-09-296-264-16
Perfect score: 20
Sequence: 1 tcctctggcttctggtagcg 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 335578 seqs, 189365133 residues

Total number of hits satisfying chosen parameters: 195614

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	13.6	68.0	51	10	US-09-794-927-41	Sequence 41, Appl
c	2	13.6	68.0	51	10	US-09-794-927-42	Sequence 42, Appl
	3	13.6	68.0	51	10	US-09-795-847-41	Sequence 41, Appl
c	4	13.6	68.0	51	10	US-09-795-847-42	Sequence 42, Appl
	5	13.6	68.0	51	10	US-09-794-743-41	Sequence 41, Appl
c	6	13.6	68.0	51	10	US-09-794-743-42	Sequence 42, Appl
	7	13.6	68.0	51	10	US-09-794-748-41	Sequence 41, Appl
c	8	13.6	68.0	51	10	US-09-794-748-42	Sequence 42, Appl
	9	13.6	68.0	51	10	US-09-794-925-41	Sequence 41, Appl
c	10	13.6	68.0	51	10	US-09-794-925-42	Sequence 42, Appl
	11	13.6	68.0	51	10	US-09-681-442-41	Sequence 41, Appl
c	12	13.6	68.0	51	10	US-09-681-442-42	Sequence 42, Appl
	13	13.2	66.0	25	10	US-09-866-108-14440	Sequence 14440, A
	14	13.2	66.0	25	10	US-09-866-108-14441	Sequence 14441, A
	15	13.2	66.0	25	10	US-09-866-108-14442	Sequence 14442, A
	16	13.2	66.0	25	10	US-09-866-108-14443	Sequence 14443, A
	17	13.2	66.0	25	10	US-09-866-108-14444	Sequence 14444, A
	18	13.2	66.0	25	10	US-09-866-108-14445	Sequence 14445, A
	19	13.2	66.0	25	10	US-09-866-108-14446	Sequence 14446, A
	20	13.2	66.0	25	10	US-09-866-108-14447	Sequence 14447, A
c	21	12.8	64.0	45	9	US-09-905-291A-146	Sequence 146, App
c	22	12.8	64.0	45	10	US-09-909-320-146	Sequence 146, App
c	23	12.8	64.0	45	10	US-09-909-088B-146	Sequence 146, App
c	24	12.8	64.0	100	10	US-09-969-373-508	Sequence 508, App
c	25	12.6	63.0	35	10	US-09-865-483-4	Sequence 4, Appli
c	26	12.6	63.0	84	10	US-09-250-611-16	Sequence 16, Appl
	27	12.4	62.0	20	10	US-09-754-167-62	Sequence 62, Appl
c	28	12.4	62.0	23	9	US-09-978-295A-348	Sequence 348, App
c	29	12.4	62.0	23	9	US-09-978-697-348	Sequence 348, App
	30	12.4	62.0	100	10	US-09-999-672-23	Sequence 23, Appl
	31	12.4	62.0	100	10	US-09-999-672-26	Sequence 26, Appl
	32	12.4	62.0	100	12	US-10-040-863-23	Sequence 23, Appl
	33	12.4	62.0	100	12	US-10-040-863-26	Sequence 26, Appl
	34	12.2	61.0	17	10	US-09-866-108-9547	Sequence 9547, Ap
	35	12.2	61.0	17	10	US-09-866-108-9548	Sequence 9548, Ap
c	36	12.2	61.0	19	10	US-09-829-549A-3	Sequence 3, Appli
	37	12.2	61.0	25	10	US-09-866-108-14439	Sequence 14439, A
	38	12.2	61.0	25	10	US-09-866-108-14448	Sequence 14448, A
	39	12.2	61.0	27	10	US-09-867-274-21	Sequence 21, Appl
c	40	12.2	61.0	36	9	US-09-966-955A-50	Sequence 50, Appl
c	41	12.2	61.0	87	9	US-09-832-659-29	Sequence 29, Appl
c	42	12.2	61.0	92	10	US-09-864-761-28968	Sequence 28968, A
c	43	12.2	61.0	99	10	US-09-969-708-195	Sequence 195, App
	44	12	60.0	18	10	US-09-969-373-3103	Sequence 3103, Ap
c	45	12	60.0	34	10	US-09-852-922-18	Sequence 18, Appl

ALIGNMENTS

RESULT 1
US-09-794-927-41
; Sequence 41, Application US/09794927
; Patent No. US20010016324A1
; GENERAL INFORMATION:
; APPLICANT: Gurney, Mark E.
; APPLICANT: Bienkowski, Michael J.
; APPLICANT: Heinrikson, Robert L.
; APPLICANT: Parodi, Luis A.
; APPLICANT: Yan, Riqiang
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AND
; TITLE OF INVENTION: USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: 28341/6280FG
; CURRENT APPLICATION NUMBER: US/09/794,927
; CURRENT FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: 09/416,901
; PRIOR FILING DATE: 1999-10-13
; PRIOR APPLICATION NUMBER: 60/155,493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 09/404,133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 60/101,594
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 41
; LENGTH: 51
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Caspase 8
; OTHER INFORMATION: Cleavage Site
US-09-794-927-41

Query Match 68.0%; Score 13.6; DB 10; Length 51;
Best Local Similarity 80.0%; Pred. No. 6.2e+02;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 TCCTCTGGCTTCTGGTAGCG 20

```
          ||| ||||  ||||||| ||
Db      23 TCCGCTGGACTCTGGTATCG 42


RESULT 2
US-09-794-927-42/c
; Sequence 42, Application US/09794927
; Patent No. US20010016324A1
; GENERAL INFORMATION:
;  APPLICANT:  Gurney, Mark E.
;  APPLICANT:  Bienkowski, Michael J.
;  APPLICANT:  Heinrikson, Robert L.
;  APPLICANT:  Parodi, Luis A.
;  APPLICANT:  Yan, Riqiang
;  TITLE OF INVENTION:  ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AND
;  TITLE OF INVENTION:  USES
;  TITLE OF INVENTION:  THEREFOR
;  FILE REFERENCE:  28341/6280FG
;  CURRENT APPLICATION NUMBER: US/09/794,927
;  CURRENT FILING DATE:  2001-02-27
;  PRIOR APPLICATION NUMBER:  09/416,901
;  PRIOR FILING DATE:  1999-10-13
;  PRIOR APPLICATION NUMBER:  60/155,493
;  PRIOR FILING DATE:  1999-09-23
;  PRIOR APPLICATION NUMBER:  09/404,133
;  PRIOR FILING DATE:  1999-09-23
;  PRIOR APPLICATION NUMBER:  PCT/US99/20881
;  PRIOR FILING DATE:  1999-09-23
;  PRIOR APPLICATION NUMBER:  60/101,594
;  PRIOR FILING DATE:  1998-09-24
;  NUMBER OF SEQ ID NOS:  73
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 42
;   LENGTH: 51
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Caspase 8
;   OTHER INFORMATION: Cleavage Site
US-09-794-927-42

  Query Match             68.0%;  Score 13.6;  DB 10;  Length 51;
  Best Local Similarity   80.0%;  Pred. No. 6.2e+02;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 TCCTCTGGCTTCTGGTAGCG 20
           ||| ||||  ||||||| ||
Db      33 TCCGCTGGACTCTGGTATCG 14


RESULT 3
US-09-795-847-41
; Sequence 41, Application US/09795847
; Patent No. US20010018208A1
; GENERAL INFORMATION:
;  APPLICANT:  Gurney, Mark E.
;  APPLICANT:  Bienkowski, Michael J.
;  APPLICANT:  Heinrikson, Robert L.
;  APPLICANT:  Parodi, Luis A.
;  APPLICANT:  Yan, Riqiang
;  TITLE OF INVENTION:  ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AND
;  TITLE OF INVENTION:  USES
;  TITLE OF INVENTION:  THEREFOR
;  FILE REFERENCE:  28341/6280DE
;  CURRENT APPLICATION NUMBER: US/09/795,847
;  CURRENT FILING DATE:  2001-02-28
;  PRIOR APPLICATION NUMBER:  09/416,901
;  PRIOR FILING DATE:  1999-10-13
;  PRIOR APPLICATION NUMBER:  60/155,493
;  PRIOR FILING DATE:  1999-09-23
;  PRIOR APPLICATION NUMBER:  09/404,133
;  PRIOR FILING DATE:  1999-09-23
;  PRIOR APPLICATION NUMBER:  PCT/US99/20881
;  PRIOR FILING DATE:  1999-09-23
;  PRIOR APPLICATION NUMBER:  60/101,594
;  PRIOR FILING DATE:  1998-09-24
;  NUMBER OF SEQ ID NOS:  73
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 41
;   LENGTH: 51
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Caspase 8
;   OTHER INFORMATION: Cleavage Site
US-09-795-847-41

  Query Match             68.0%;  Score 13.6;  DB 10;  Length 51;
  Best Local Similarity   80.0%;  Pred. No. 6.2e+02;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 TCCTCTGGCTTCTGGTAGCG 20
           ||| ||||  ||||||| ||
Db      23 TCCGCTGGACTCTGGTATCG 42


RESULT 4
US-09-795-847-42/c
; Sequence 42, Application US/09795847
; Patent No. US20010018208A1
; GENERAL INFORMATION:
;  APPLICANT:  Gurney, Mark E.
;  APPLICANT:  Bienkowski, Michael J.
;  APPLICANT:  Heinrikson, Robert L.
;  APPLICANT:  Parodi, Luis A.
;  APPLICANT:  Yan, Riqiang
;  TITLE OF INVENTION:  ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AND
;  TITLE OF INVENTION:  USES
;  TITLE OF INVENTION:  THEREFOR
;  FILE REFERENCE:  28341/6280DE
;  CURRENT APPLICATION NUMBER: US/09/795,847
;  CURRENT FILING DATE:  2001-02-28
;  PRIOR APPLICATION NUMBER:  09/416,901
;  PRIOR FILING DATE:  1999-10-13
;  PRIOR APPLICATION NUMBER:  60/155,493
;  PRIOR FILING DATE:  1999-09-23
;  PRIOR APPLICATION NUMBER:  09/404,133
;  PRIOR FILING DATE:  1999-09-23
;  PRIOR APPLICATION NUMBER:  PCT/US99/20881
;  PRIOR FILING DATE:  1999-09-23
;  PRIOR APPLICATION NUMBER:  60/101,594
;  PRIOR FILING DATE:  1998-09-24
;  NUMBER OF SEQ ID NOS:  73
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 42
;   LENGTH: 51
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Caspase 8
;   OTHER INFORMATION: Cleavage Site
US-09-795-847-42

  Query Match             68.0%;  Score 13.6;  DB 10;  Length 51;
  Best Local Similarity   80.0%;  Pred. No. 6.2e+02;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 TCCTCTGGCTTCTGGTAGCG 20
           ||| ||||  ||||||| ||
Db      33 TCCGCTGGACTCTGGTATCG 14


RESULT 5
US-09-794-743-41
; Sequence 41, Application US/09794743
; Patent No. US20010021391A1
```

```
; GENERAL INFORMATION:
;  APPLICANT:  Gurney, Mark E.
;  APPLICANT:  Bienkowski, Michael J.
;  APPLICANT:  Heinrikson, Robert L.
;  APPLICANT:  Parodi, Luis A.
;  APPLICANT:  Yan, Riqiang
;  TITLE OF INVENTION:  ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AND
;  TITLE OF INVENTION:  USES
;  TITLE OF INVENTION:  THEREFOR
;  FILE REFERENCE:  28341/6280BC
;  CURRENT APPLICATION NUMBER: US/09/794,743
;  CURRENT FILING DATE:  2001-02-27
;  PRIOR APPLICATION NUMBER:  09/416,901
;  PRIOR FILING DATE:  1999-10-13
;  PRIOR APPLICATION NUMBER:  60/155,493
;  PRIOR FILING DATE:  1999-09-23
;  PRIOR APPLICATION NUMBER:  09/404,133
;  PRIOR FILING DATE:  1999-09-23
;  PRIOR APPLICATION NUMBER:  PCT/US99/20881
;  PRIOR FILING DATE:  1999-09-23
;  PRIOR APPLICATION NUMBER:  60/101,594
;  PRIOR FILING DATE:  1998-09-24
;  NUMBER OF SEQ ID NOS:  73
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 41
;   LENGTH: 51
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Caspase 8
;   OTHER INFORMATION: Cleavage Site
US-09-794-743-41

  Query Match             68.0%;  Score 13.6;  DB 10;  Length 51;
  Best Local Similarity   80.0%;  Pred. No. 6.2e+02;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 TCCTCTGGCTTCTGGTAGCG 20
           ||| ||||  ||||||| ||
Db      23 TCCGCTGGACTCTGGTATCG 42


RESULT 6
US-09-794-743-42/c
; Sequence 42, Application US/09794743
; Patent No. US20010021391A1
; GENERAL INFORMATION:
;  APPLICANT:  Gurney, Mark E.
;  APPLICANT:  Bienkowski, Michael J.
;  APPLICANT:  Heinrikson, Robert L.
;  APPLICANT:  Parodi, Luis A.
;  APPLICANT:  Yan, Riqiang
;  TITLE OF INVENTION:  ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AND
;  TITLE OF INVENTION:  USES
;  TITLE OF INVENTION:  THEREFOR
;  FILE REFERENCE:  28341/6280BC
;  CURRENT APPLICATION NUMBER: US/09/794,743
;  CURRENT FILING DATE:  2001-02-27
;  PRIOR APPLICATION NUMBER:  09/416,901
;  PRIOR FILING DATE:  1999-10-13
;  PRIOR APPLICATION NUMBER:  60/155,493
;  PRIOR FILING DATE:  1999-09-23
;  PRIOR APPLICATION NUMBER:  09/404,133
;  PRIOR FILING DATE:  1999-09-23
;  PRIOR APPLICATION NUMBER:  PCT/US99/20881
;  PRIOR FILING DATE:  1999-09-23
;  PRIOR APPLICATION NUMBER:  60/101,594
;  PRIOR FILING DATE:  1998-09-24
;  NUMBER OF SEQ ID NOS:  73
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 42
;   LENGTH: 51
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Caspase 8
;   OTHER INFORMATION: Cleavage Site
US-09-794-743-42

  Query Match             68.0%;  Score 13.6;  DB 10;  Length 51;
  Best Local Similarity   80.0%;  Pred. No. 6.2e+02;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 TCCTCTGGCTTCTGGTAGCG 20
           ||| ||||  ||||||| ||
Db      33 TCCGCTGGACTCTGGTATCG 14


RESULT 7
US-09-794-748-41
; Sequence 41, Application US/09794748
; Patent No. US20020037315A1
; GENERAL INFORMATION:
;  APPLICANT:  Gurney, Mark E.
;  APPLICANT:  Bienkowski, Michael J.
;  APPLICANT:  Heinrikson, Robert L.
;  APPLICANT:  Parodi, Luis A.
;  APPLICANT:  Yan, Riqiang
;  TITLE OF INVENTION:  ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AND
;  TITLE OF INVENTION:  USES
;  TITLE OF INVENTION:  THEREFOR
;  FILE REFERENCE:  28341/6280JL
;  CURRENT APPLICATION NUMBER: US/09/794,748
;  CURRENT FILING DATE:  2001-02-27
;  PRIOR APPLICATION NUMBER:  09/416,901
;  PRIOR FILING DATE:  1999-10-13
;  PRIOR APPLICATION NUMBER:  60/155,493
;  PRIOR FILING DATE:  1999-09-23
;  PRIOR APPLICATION NUMBER:  09/404,133
;  PRIOR FILING DATE:  1999-09-23
;  PRIOR APPLICATION NUMBER:  PCT/US99/20881
;  PRIOR FILING DATE:  1999-09-23
;  PRIOR APPLICATION NUMBER:  60/101,594
;  PRIOR FILING DATE:  1998-09-24
;  NUMBER OF SEQ ID NOS:  73
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 41
;   LENGTH: 51
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Caspase 8
;   OTHER INFORMATION: Cleavage Site
US-09-794-748-41

  Query Match             68.0%;  Score 13.6;  DB 10;  Length 51;
  Best Local Similarity   80.0%;  Pred. No. 6.2e+02;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 TCCTCTGGCTTCTGGTAGCG 20
           ||| ||||  ||||||| ||
Db      23 TCCGCTGGACTCTGGTATCG 42


RESULT 8
US-09-794-748-42/c
; Sequence 42, Application US/09794748
; Patent No. US20020037315A1
; GENERAL INFORMATION:
;  APPLICANT:  Gurney, Mark E.
;  APPLICANT:  Bienkowski, Michael J.
;  APPLICANT:  Heinrikson, Robert L.
;  APPLICANT:  Parodi, Luis A.
;  APPLICANT:  Yan, Riqiang
;  TITLE OF INVENTION:  ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AND
;  TITLE OF INVENTION:  USES
```

```
;  TITLE OF INVENTION:  THEREFOR
;  FILE REFERENCE:  28341/6280JL
;  CURRENT APPLICATION NUMBER: US/09/794,748
;  CURRENT FILING DATE:  2001-02-27
;  PRIOR APPLICATION NUMBER:  09/416,901
;  PRIOR FILING DATE:  1999-10-13
;  PRIOR APPLICATION NUMBER:  60/155,493
;  PRIOR FILING DATE:  1999-09-23
;  PRIOR APPLICATION NUMBER:  09/404,133
;  PRIOR FILING DATE:  1999-09-23
;  PRIOR APPLICATION NUMBER:  PCT/US99/20881
;  PRIOR FILING DATE:  1999-09-23
;  PRIOR APPLICATION NUMBER:  60/101,594
;  PRIOR FILING DATE:  1998-09-24
;  NUMBER OF SEQ ID NOS:  73
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 42
;   LENGTH: 51
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Caspase 8
;   OTHER INFORMATION: Cleavage Site
US-09-794-748-42

  Query Match             68.0%;  Score 13.6;  DB 10;  Length 51;
  Best Local Similarity   80.0%;  Pred. No. 6.2e+02;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 TCCTCTGGCTTCTGGTAGCG 20
           ||| ||||  ||||||| ||
Db      33 TCCGCTGGACTCTGGTATCG 14


RESULT 9
US-09-794-925-41
; Sequence 41, Application US/09794925
; Patent No. US20020064819A1
; GENERAL INFORMATION:
;  APPLICANT:  Gurney, Mark E.
;  APPLICANT:  Bienkowski, Michael J.
;  APPLICANT:  Heinrikson, Robert L.
;  APPLICANT:  Parodi, Luis A.
;  APPLICANT:  Yan, Riqiang
;  TITLE OF INVENTION:  ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AND USES
;  TITLE OF INVENTION:  THEREFOR
;  FILE REFERENCE:  28341/6280HI
;  CURRENT APPLICATION NUMBER: US/09/794,925
;  CURRENT FILING DATE:  2001-02-27
;  PRIOR APPLICATION NUMBER:  09/416,901
;  PRIOR FILING DATE:  1999-10-13
;  PRIOR APPLICATION NUMBER:  60/155,493
;  PRIOR FILING DATE:  1999-09-23
;  PRIOR APPLICATION NUMBER:  09/404,133
;  PRIOR FILING DATE:  1999-09-23
;  PRIOR APPLICATION NUMBER:  PCT/US99/20881
;  PRIOR FILING DATE:  1999-09-23
;  PRIOR APPLICATION NUMBER:  60/101,594
;  PRIOR FILING DATE:  1998-09-24
;  NUMBER OF SEQ ID NOS:  73
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 41
;   LENGTH: 51
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Caspase 8
;   OTHER INFORMATION: Cleavage Site
US-09-794-925-41

  Query Match             68.0%;  Score 13.6;  DB 10;  Length 51;
  Best Local Similarity   80.0%;  Pred. No. 6.2e+02;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 TCCTCTGGCTTCTGGTAGCG 20
           ||| ||||  ||||||| ||
Db      23 TCCGCTGGACTCTGGTATCG 42


RESULT 10
US-09-794-925-42/c
; Sequence 42, Application US/09794925
; Patent No. US20020064819A1
; GENERAL INFORMATION:
;  APPLICANT:  Gurney, Mark E.
;  APPLICANT:  Bienkowski, Michael J.
;  APPLICANT:  Heinrikson, Robert L.
;  APPLICANT:  Parodi, Luis A.
;  APPLICANT:  Yan, Riqiang
;  TITLE OF INVENTION:  ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AND U
;  TITLE OF INVENTION:  THEREFOR
;  FILE REFERENCE:  28341/6280HI
;  CURRENT APPLICATION NUMBER: US/09/794,925
;  CURRENT FILING DATE:  2001-02-27
;  PRIOR APPLICATION NUMBER:  09/416,901
;  PRIOR FILING DATE:  1999-10-13
;  PRIOR APPLICATION NUMBER:  60/155,493
;  PRIOR FILING DATE:  1999-09-23
;  PRIOR APPLICATION NUMBER:  09/404,133
;  PRIOR FILING DATE:  1999-09-23
;  PRIOR APPLICATION NUMBER:  PCT/US99/20881
;  PRIOR FILING DATE:  1999-09-23
;  PRIOR APPLICATION NUMBER:  60/101,594
;  PRIOR FILING DATE:  1998-09-24
;  NUMBER OF SEQ ID NOS:  73
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 42
;   LENGTH: 51
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Caspase 8
;   OTHER INFORMATION: Cleavage Site
US-09-794-925-42

  Query Match             68.0%;  Score 13.6;  DB 10;  Length 51;
  Best Local Similarity   80.0%;  Pred. No. 6.2e+02;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 TCCTCTGGCTTCTGGTAGCG 20
           ||| ||||  ||||||| ||
Db      33 TCCGCTGGACTCTGGTATCG 14


RESULT 11
US-09-681-442-41
; Sequence 41, Application US/09681442
; Patent No. US20020081634A1
; GENERAL INFORMATION:
;  APPLICANT: Gurney, Mark E.
;  APPLICANT:  Bienkowski, Michael J.
;  APPLICANT:  Heinrikson, Robert L.
;  APPLICANT:  Parodi, Luis A.
;  APPLICANT:  Yan, Riqiang
;  TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AND US
;  TITLE OF INVENTION:  THEREFOR
;  FILE REFERENCE: 28341/6280FG
;  CURRENT APPLICATION NUMBER: US/09/681,442
;  CURRENT FILING DATE:  2001-04-05
;  PRIOR APPLICATION NUMBER: 09/416,901
;  PRIOR FILING DATE: 1999-10-13
;  PRIOR APPLICATION NUMBER: 60/155,493
;  PRIOR FILING DATE: 1999-09-23
;  PRIOR APPLICATION NUMBER: 09/404,133
;  PRIOR FILING DATE: 1999-09-23
;  PRIOR APPLICATION NUMBER: PCT/US99/20881
```

```
;  PRIOR FILING DATE: 1999-09-23
;  PRIOR APPLICATION NUMBER: 60/101,594
;  PRIOR FILING DATE: 1998-09-24
;  NUMBER OF SEQ ID NOS: 73
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 41
;   LENGTH: 51
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Caspase 8
;   OTHER INFORMATION: Cleavage Site
US-09-681-442-41

  Query Match             68.0%;  Score 13.6;  DB 10;  Length 51;
  Best Local Similarity   80.0%;  Pred. No. 6.2e+02;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 TCCTCTGGCTTCTGGTAGCG 20
           ||| ||||  ||||||| ||
Db      23 TCCGCTGGACTCTGGTATCG 42


RESULT 12
US-09-681-442-42/c
; Sequence 42, Application US/09681442
; Patent No. US20020081634A1
; GENERAL INFORMATION:
;  APPLICANT: Gurney, Mark E.
;  APPLICANT:  Bienkowski, Michael J.
;  APPLICANT:  Heinrikson, Robert L.
;  APPLICANT:  Parodi, Luis A.
;  APPLICANT:  Yan, Riqiang
;  TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AND USES
;  TITLE OF INVENTION:  THEREFOR
;  FILE REFERENCE: 28341/6280FG
;  CURRENT APPLICATION NUMBER: US/09/681,442
;  CURRENT FILING DATE:  2001-04-05
;  PRIOR APPLICATION NUMBER: 09/416,901
;  PRIOR FILING DATE: 1999-10-13
;  PRIOR APPLICATION NUMBER: 60/155,493
;  PRIOR FILING DATE: 1999-09-23
;  PRIOR APPLICATION NUMBER: 09/404,133
;  PRIOR FILING DATE: 1999-09-23
;  PRIOR APPLICATION NUMBER: PCT/US99/20881
;  PRIOR FILING DATE: 1999-09-23
;  PRIOR APPLICATION NUMBER: 60/101,594
;  PRIOR FILING DATE: 1998-09-24
;  NUMBER OF SEQ ID NOS: 73
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 42
;   LENGTH: 51
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Caspase 8
;   OTHER INFORMATION: Cleavage Site
US-09-681-442-42

  Query Match             68.0%;  Score 13.6;  DB 10;  Length 51;
  Best Local Similarity   80.0%;  Pred. No. 6.2e+02;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 TCCTCTGGCTTCTGGTAGCG 20
           ||| ||||  ||||||| ||
Db      33 TCCGCTGGACTCTGGTATCG 14


RESULT 13
US-09-866-108-14440
; Sequence 14440, Application US/09866108
; Patent No. US20020048800A1
; GENERAL INFORMATION:
;  APPLICANT: GU, Yizhong
;  APPLICANT:  JI, Yonggang
;  APPLICANT:  PENN, Sharron G.
;  APPLICANT:  HANZEL, David K.
;  APPLICANT:  RANK, David R.
;  APPLICANT:  CHEN, Wensheng
;  APPLICANT:  SHANNON, Mark
;  TITLE OF INVENTION: MYOSIN-LIKE GENE EXPRESSED IN HUMAN HEART AND MUSCLE
;  FILE REFERENCE: AEOMICA-7
;  CURRENT APPLICATION NUMBER: US/09/866,108
;  CURRENT FILING DATE:  2001-05-25
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 60/266,860
;  PRIOR FILING DATE: 2001-02-05
;  NUMBER OF SEQ ID NOS: 15752
;  SOFTWARE: Aeomica Sequence Listing Engine
; SEQ ID NO 14440
;   LENGTH: 25
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-866-108-14440

  Query Match             66.0%;  Score 13.2;  DB 10;  Length 25;
  Best Local Similarity   83.3%;  Pred. No. 8.9e+02;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 TCCTCTGGCTTCTGGTAG 18
           ||| ||||||||| | ||
Db       8 TCCCCTGGCTTCTCGGAG 25


RESULT 14
US-09-866-108-14441
; Sequence 14441, Application US/09866108
; Patent No. US20020048800A1
; GENERAL INFORMATION:
;  APPLICANT: GU, Yizhong
;  APPLICANT:  JI, Yonggang
;  APPLICANT:  PENN, Sharron G.
;  APPLICANT:  HANZEL, David K.
;  APPLICANT:  RANK, David R.
;  APPLICANT:  CHEN, Wensheng
;  APPLICANT:  SHANNON, Mark
;  TITLE OF INVENTION: MYOSIN-LIKE GENE EXPRESSED IN HUMAN HEART AND MUSCLE
;  FILE REFERENCE: AEOMICA-7
;  CURRENT APPLICATION NUMBER: US/09/866,108
```

```
;  CURRENT FILING DATE:  2001-05-25
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 60/266,860
;  PRIOR FILING DATE: 2001-02-05
;  NUMBER OF SEQ ID NOS: 15752
;  SOFTWARE: Aeomica Sequence Listing Engine
; SEQ ID NO 14441
;   LENGTH: 25
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-866-108-14441

  Query Match             66.0%;  Score 13.2;  DB 10;  Length 25;
  Best Local Similarity   83.3%;  Pred. No. 8.9e+02;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 TCCTCTGGCTTCTGGTAG 18
          ||| |||||||||  | ||
Db      7 TCCCCTGGCTTCTCGGAG 24


RESULT 15
US-09-866-108-14442
; Sequence 14442, Application US/09866108
; Patent No. US20020048800A1
; GENERAL INFORMATION:
;  APPLICANT: GU, Yizhong
;  APPLICANT:  JI, Yonggang
;  APPLICANT:  PENN, Sharron G.
;  APPLICANT:  HANZEL, David K.
;  APPLICANT:  RANK, David R.
;  APPLICANT:  CHEN, Wensheng
;  APPLICANT:  SHANNON, Mark
;  TITLE OF INVENTION: MYOSIN-LIKE GENE EXPRESSED IN HUMAN HEART AND MUSCLE
;  FILE REFERENCE: AEOMICA-7
;  CURRENT APPLICATION NUMBER: US/09/866,108
;  CURRENT FILING DATE:  2001-05-25
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 60/266,860
;  PRIOR FILING DATE: 2001-02-05
;  NUMBER OF SEQ ID NOS: 15752
;  SOFTWARE: Aeomica Sequence Listing Engine
; SEQ ID NO 14442
;   LENGTH: 25
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-866-108-14442

  Query Match             66.0%;  Score 13.2;  DB 10;  Length 25;
  Best Local Similarity   83.3%;  Pred. No. 8.9e+02;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 TCCTCTGGCTTCTGGTAG 18
          ||| |||||||||  | ||
Db      6 TCCCCTGGCTTCTCGGAG 23


Search completed: November 23, 2002, 06:42:12
Job time : 17.25 secs
```

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 25, 2002, 09:10:06 ; Search time 755.55 Seconds
(without alignments)
428.707 Million cell updates/sec

Title: US-09-296-264-16
Perfect score: 20
Sequence: 1 tcctctggcttctggtagcg 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 357874

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	14.4	72.0	32	10	AV833085	AV833085 AV833085
c	2	14.2	71.0	52	9	AI528400	AI528400 ui97a04.y
	3	14.2	71.0	59	12	BG342279	BG342279 602374371
	4	14.2	71.0	70	9	AA470781	AA470781 ne19e12.s
	5	14.2	71.0	82	9	AA937053	AA937053 ol87a07.s
	6	14.2	71.0	89	9	AA804881	AA804881 ob99a03.s
	7	13.8	69.0	79	12	BG112276	BG112276 602282420
	8	13.8	69.0	82	17	BH228133	BH228133 1006144D1
	9	13.8	69.0	100	17	AZ424155	AZ424155 1M0203C08
c	10	13.6	68.0	70	9	AI430606	AI430606 me08a02.y
c	11	13.4	67.0	52	9	AA894524	AA894524 of90c08.s
	12	13.4	67.0	75	9	AU258999	AU258999 AU258999
c	13	13.4	67.0	94	9	AI120925	AI120925 ub74d04.r
c	14	13.4	67.0	96	12	BG550793	BG550793 sad24a03.
c	15	13.4	67.0	99	17	AZ820877	AZ820877 2M0093G21
c	16	13.2	66.0	53	14	BM893982	BM893982 ij33e08.x
c	17	13.2	66.0	60	13	BJ034239	BJ034239 BJ034239
c	18	13.2	66.0	68	13	BI194584	BI194584 602946463
	19	13.2	66.0	82	9	AI021531	AI021531 ub09b03.r
c	20	13.2	66.0	85	9	AA564543	AA564543 nj02b10.s
	21	13.2	66.0	86	14	BM888226	BM888226 TMM142 Hu
c	22	13.2	66.0	86	17	CNS041PL	AL270498 Tetraodon
	23	13.2	66.0	87	13	BI090121	BI090121 602854845
c	24	13.2	66.0	89	9	AA640184	AA640184 np28e09.s
	25	13.2	66.0	90	9	AA200557	AA200557 mu34f01.r
	26	13.2	66.0	99	9	AA561525	AA561525 vl28a10.r
	27	13.2	66.0	99	14	BM967335	BM967335 ij33e08.y
c	28	12.8	64.0	61	14	BQ099919	BQ099919 ph70b07.y
c	29	12.8	64.0	63	9	AI748404	AI748404 sb52d06.y
c	30	12.8	64.0	78	13	BI191352	BI191352 k2f01fs.r
c	31	12.8	64.0	79	17	AZ645149	AZ645149 1M0510K21
c	32	12.8	64.0	83	17	AZ308826	AZ308826 1M0012B13
	33	12.8	64.0	88	17	AZ430802	AZ430802 1M0215M19
	34	12.8	64.0	94	17	AZ778982	AZ778982 2M0014A15
	35	12.8	64.0	99	9	AU133902	AU133902 AU133902
	36	12.6	63.0	42	17	L48765	L48765 HUMBA Chrom
c	37	12.6	63.0	42	17	TA131E09P	AL464188 T. brucei
c	38	12.6	63.0	51	10	AW643833	AW643833 cm34d06.w
	39	12.6	63.0	51	17	AZ852266	AZ852266 2M0154J23
	40	12.6	63.0	53	12	BE906112	BE906112 601497159
c	41	12.6	63.0	58	9	AA518037	AA518037 vi02b05.r
c	42	12.6	63.0	61	9	AA509513	AA509513 vg19e11.r
	43	12.6	63.0	76	12	BG083136	BG083136 H3084H06-
c	44	12.6	63.0	78	9	AA783070	AA783070 a1h06f2.f
	45	12.6	63.0	80	17	AZ631141	AZ631141 1M0485J11

ALIGNMENTS

RESULT 1
AV833085
LOCUS AV833085 32 bp mRNA linear EST 22-JUN-2001
DEFINITION AV833085 K. Sato unpublished cDNA library: Hordeum vulgare subsp. vulgare shoots germination Hordeum vulgare subsp. vulgare cDNA clone bags11h20, mRNA sequence.
ACCESSION AV833085
VERSION AV833085.1 GI:14525174
KEYWORDS EST.
SOURCE Hordeum vulgare subsp. vulgare.
ORGANISM Hordeum vulgare subsp. vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae ; Triticeae; Hordeum.
REFERENCE 1 (bases 1 to 32)
AUTHORS Sato,K.
TITLE Barley EST sequencing project in NIG and Okayama Univ
JOURNAL Unpublished (2001)
COMMENT Contact: Kazuhiro Sato
Research Institute for Bioresources
Okayama University, Barley Germplasm Center
Chuo 2-20-1, Kurashiki, Okayama 710-0046, Japan
Email: kazsato@rib.okayama-u.ac.jp,
URL:http://www.rib.okayama-u.ac.jp/barley/
Sato,K., Saisho,D., Takeda,K., Shini,T. and Kohara,Y. Direct submission;
database:http://www.shigen.nig.ac.jp/barley/Barley.html.
FEATURES Location/Qualifiers
source 1. .32

```
                     /organism="Hordeum vulgare subsp. vulgare"
                     /cultivar="Haruna Nijo"
                     /db_xref="taxon:112509"
                     /clone="bags11h20"
                     /clone_lib="K. Sato unpublished cDNA library: Hordeum
                     vulgare subsp. vulgare shoots germination"
                     /tissue_type="shoots"
                     /dev_stage="germination"
BASE COUNT        9 a      9 c      5 g      9 t
ORIGIN

  Query Match            72.0%;  Score 14.4;  DB 10;  Length 32;
  Best Local Similarity   93.8%;  Pred. No. 1.7e+04;
  Matches   15;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 CCTCTGGCTTCTGGTA 17
           ||||||||||||||| |
Db      13 CCTCTGGCTTCTGGAA 28


RESULT 2
AI528400/c
LOCUS       AI528400                  52 bp    mRNA    linear   EST 18-MAR-1999
DEFINITION  ui97a04.y1 Sugano mouse liver mlia Mus musculus cDNA clone
            IMAGE:1890318 5' similar to SW:LCAT_MOUSE P16301
            PHOSPHATIDYLCHOLINE-STEROL ACYLTRANSFERASE PRECURSOR ;, mRNA
            sequence.
ACCESSION   AI528400
VERSION     AI528400.1  GI:4442535
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 52)
  AUTHORS   Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,
            Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person
            ,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter
            ,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R.,
            Waterston,R. and Wilson,R.
  TITLE     The WashU-NCI Mouse EST Project 1999
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Marra M/WashU-NCI Mouse EST Project 1999
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:974642
            Trace considered overall poor quality
            Possible reversed clone: similarity on wrong strand
            Seq primer: custom primer used
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .52
                     /organism="Mus musculus"
                     /strain="C57BL"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:1890318"
                     /clone_lib="Sugano mouse liver mlia"
                     /sex="female"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /note="Organ: liver; Vector: pME18S-FL3; Site_1: DraIII
                     (CACTGTGTG); Site_2: DraIII (CACCATGTG); 1st strand cDNA
                     was primed with an oligo(dT) primer
                     [ATGTGGCCTTTTTTTTTTTTTTTT]; double-stranded cDNA was
                     ligated to a  DraIII adaptor [TGTTGGCCTACTGG], digested
                     and cloned into distinct DraIII sites of the pME18S-FL3
                     vector (5' site CACTGTGTG, 3' site CACCATGTG). XhoI should
                     be used to isolate the cDNA insert.  Size selection was
                     performed to exclude fragments <1.5kb.  Library
                     constructed by Dr. Sumio Sugano (University of Tokyo
                     Institute of Medical Science).  Custom primers  for
                     sequencing: 5' end primer CTTCTGCTCTAAAAGCTGCG and 3' end
                     primer CGACCTGCAGCTCGAGCACA."
BASE COUNT       18 a     11 c     11 g     12 t
ORIGIN

  Query Match            71.0%;  Score 14.2;  DB 9;  Length 52;
  Best Local Similarity   84.2%;  Pred. No. 2.6e+04;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 TCCTCTGGCTTCTGGTAGC 19
           ||||||| |||  ||||||
Db      39 TCCTCTGTCTTGCGGTAGC 21


RESULT 3
BG342279
LOCUS       BG342279                  59 bp    mRNA    linear   EST 27-FEB-2001
DEFINITION  602374371F1 NIH_MGC_94 Mus musculus cDNA clone IMAGE:4481872 5',
            mRNA sequence.
ACCESSION   BG342279
VERSION     BG342279.1  GI:13148717
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 59)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: The Cepko Laboratory
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10317  row: n  column: 17
            High quality sequence stop: 57.
FEATURES             Location/Qualifiers
     source          1. .59
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:4481872"
                     /clone_lib="NIH_MGC_94"
                     /tissue_type="retina"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: eye; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                     Average insert size 3.3 kb. Library enriched for
                     full-length clones and constructed by Life Technologies.
                     Note: this is a NIH_MGC Library."
BASE COUNT       18 a     15 c     11 g     15 t
ORIGIN

  Query Match            71.0%;  Score 14.2;  DB 12;  Length 59;
  Best Local Similarity   84.2%;  Pred. No. 2.7e+04;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 CCTCTGGCTTCTGGTAGCG 20
           |||||  |||||||||| |
Db      15 CCTCTTCCTTCTGGTAGTG 33


RESULT 4
AA470781
LOCUS       AA470781                  70 bp    mRNA    linear   EST 14-AUG-1997
DEFINITION  ne19e12.s1 NCI_CGAP_Co3 Homo sapiens cDNA clone IMAGE:881710 3'
```

```
            similar to SW:YBEB_ECOLI P05848 HYPOTHETICAL 7.7 KD PROTEIN IN
            MRDA/PBPA 3'REGION. ;, mRNA sequence.
ACCESSION   AA470781
VERSION     AA470781.1  GI:2198090
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 70)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Elias Campo, M.D., Michael R. Emmert-Buck, M.D.
            , Ph.D.
            cDNA Library Preparation: M. Bento Soares, Ph.D.
            cDNA Library Arrayed by: Greg Lennon, Ph.D.
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 1080   Std Error: 0.00
            Seq primer: -41m13 fwd. ET from Amersham
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .70
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:881710"
                     /clone_lib="NCI_CGAP_Co3"
                     /sex="pooled"
                     /tissue_type="colon"
                     /lab_host="DH10B"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                     was prepared from 12 pooled bulk tumor samples and primed
                     with a Not I - oligo(dT) primer. Double-stranded cDNA was
                     ligated to Eco RI adaptors (Pharmacia), digested with Not
                     I and cloned into the Not I and Eco RI sites of the
                     modified pT7T3 vector. Library went through one round of
                     normalization. "
BASE COUNT       20 a     17 c     10 g     23 t
ORIGIN

  Query Match             71.0%;  Score 14.2;  DB 9;  Length 70;
  Best Local Similarity   84.2%;  Pred. No. 2.9e+04;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 TCCTCTGGCTTCTGGTAGC 19
           | |||||| |||||| |||
Db      26 TTCTCTGGTTTCTGGAAGC 44


RESULT 5
AA937053
LOCUS       AA937053                 82 bp    mRNA    linear   EST 26-AUG-1998
DEFINITION  ol87a07.s1 NCI_CGAP_Kid5 Homo sapiens cDNA clone IMAGE:1536564 3'
            similar to TR:P73658 P73658 PUTATIVE IOJAP PROTEIN. ;, mRNA
            sequence.
ACCESSION   AA937053
VERSION     AA937053.1  GI:3094674
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 82)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
            cDNA Library Preparation: M. Bento Soares, Ph.D.
            cDNA Library Arrayed by: Greg Lennon, Ph.D.
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html

            Trace considered overall poor quality
            Insert Length: 795   Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .82
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1536564"
                     /clone_lib="NCI_CGAP_Kid5"
                     /tissue_type="2 pooled tumors (clear cell type)"
                     /lab_host="DH10B"
                     /note="Organ: kidney; Vector: pT7T3D-Pac (Pharmacia) with
                     a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer [5'
                     AACTGGAAGAATTCGCGGCCGCAATATTTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. Library
                     went through one round of normalization. Library
                     constructed by Bento Soares and M. Fatima Bonaldo.  "
BASE COUNT    .  31 a     21 c     13 g     17 t
ORIGIN

  Query Match             71.0%;  Score 14.2;  DB 9;  Length 82;
  Best Local Similarity   84.2%;  Pred. No. 3.1e+04;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 TCCTCTGGCTTCTGGTAGC 19
           | |||||| |||||| |||
Db      23 TTCTCTGGGTTCTGGAAGC 41


RESULT 6
AA804881
LOCUS       AA804881                 89 bp    mRNA    linear   EST 19-FEB-1998
DEFINITION  ob99a03.s1 NCI_CGAP_GCB1 Homo sapiens cDNA clone IMAGE:1339468 3',
            mRNA sequence.
ACCESSION   AA804881
VERSION     AA804881.1  GI:2876208
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 89)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman,
            Ph.D., Gerald Marti, M.D.
            cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
            Bonaldo, Ph.D.
            cDNA Library Arrayed by: Greg Lennon, Ph.D.
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
```

```
           Insert Length: 847   Std Error: 0.00
           Seq primer: -40m13 fwd. ET from Amersham.
FEATURES             Location/Qualifiers
     source          1. .89
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1339468"
                     /clone_lib="NCI_CGAP_GCB1"
                     /tissue_type="germinal center B cell"
                     /lab_host="DH10B"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                     was prepared from human tonsillar cells enriched for
                     germinal center B cells by flow sorting (CD20+, IgD-),
                     provided by Dr. Louis M. Staudt (NCI), Dr. David Allman
                     (NCI) and Dr. Gerald Marti (CBER).  cDNA synthesis was
                     primed with a Not I - oligo(dT) primer
                     [5'-TGTTACCAATCTGAAGTGGGAGCGGCCGCCTCATTTTTTTTTTTTTTTTTT-3'
                     ]. Double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. Library
                     went through one round of normalization, and was
                     constructed by Bento Soares and M. Fatima Bonaldo."
BASE COUNT       28 a     20 c     14 g     27 t
ORIGIN

  Query Match             71.0%;  Score 14.2;  DB 9;  Length 89;
  Best Local Similarity   84.2%;  Pred. No. 3.2e+04;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 TCCTCTGGCTTCTGGTAGC 19
           | |||||| |||||| |||
Db      41 TTCTCTGGTTTCTGGAAGC 59


RESULT 7
BG112276
LOCUS       BG112276                 79 bp    mRNA    linear   EST 30-JAN-2001
DEFINITION  602282420F1 NIH_MGC_86 Homo sapiens cDNA clone IMAGE:4369949 5',
            mRNA sequence.
ACCESSION   BG112276
VERSION     BG112276.1  GI:12605782
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 79)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10026  row: g  column: 06
            High quality sequence stop: 79.
FEATURES             Location/Qualifiers
     source          1. .79
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4369949"
                     /clone_lib="NIH_MGC_86"
                     /tissue_type="osteosarcoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: bone; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                     Average insert size 1.533 kb. Library enriched for
                     full-length clones and constructed by Life Technologies.
                     Note: this is a NIH_MGC Library."
BASE COUNT        6 a     23 c     33 g     17 t
ORIGIN

  Query Match             69.0%;  Score 13.8;  DB 12;  Length 79;
  Best Local Similarity   88.2%;  Pred. No. 4.5e+04;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 TCCTCTGGCTTCTGGTA 17
           |||||||||||||| | |
Db      22 TCCTCTGGCTTCTTGGA 38


RESULT 8
BH228133
LOCUS       BH228133                 82 bp    DNA     linear   GSS 08-NOV-2001
DEFINITION  1006144D12.x1 1006 - RescueMu Grid G Zea mays genomic, DNA
            sequence.
ACCESSION   BH228133
VERSION     BH228133.1  GI:16828846
KEYWORDS    GSS.
SOURCE      Zea mays.
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 82)
  AUTHORS   Walbot,V.
  TITLE     Maize genomic sequences found using engineered RescueMu transposon
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Walbot V
            Department of Biological Sciences
            Stanford University
            855 California Ave, Palo Alto, CA 94304, USA
            Tel: 650 723 2227
            Fax: 650 725 8221
            Email: walbot@stanford.edu
            Very probable ligation site of ends cut by single endonuclease.
            Reverse complemented post-ligation sequence from source sequence.
            Plate: 1006144  row: 5
            Class: transposon-tagged.
FEATURES             Location/Qualifiers
     source          1. .82
                     /organism="Zea mays"
                     /cultivar="mixed background W23/A188/B73"
                     /db_xref="taxon:4577"
                     /clone_lib="1006 - RescueMu Grid G"
                     /tissue_type="leaf"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /note="Organ: leaf; Vector: RescueMu (engineered from
                     pBlueScript backbone); Site_1: BamHI; Site_2: BglII;
                     RescueMu is a 4.9 kb, modified maize Mu transposon
                     designed to allow plasmid rescue from total genomic DNA.
                     Mu elements insert preferentially into transcription
                     units. For more information on RescueMu, go to the web
                     site 'www.zmdb.iastate.edu' and follow the links for
                     'RescueMu.' Grid G was grown at Stanford in 2000. DNA was
                     extracted from leaf punches, double digested using BamHI
                     and BglII, and ligated to form circular plasmids. DH10B
                     cells were transformed and then screened on LB plates with
                     ampicillin."
BASE COUNT       12 a     23 c     20 g     26 t      1 others
ORIGIN

  Query Match             69.0%;  Score 13.8;  DB 17;  Length 82;
  Best Local Similarity   88.2%;  Pred. No. 4.6e+04;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       3 CTCTGGCTTCTGGTAGC 19
           |||||| |||||| |||
Db      37 CTCTGGATTCTGGAAGC 53
```

```
RESULT 9
AZ424155
LOCUS       AZ424155                 100 bp    DNA     linear   GSS 03-OCT-2000
DEFINITION  1M0203C08R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0203C08 R, DNA sequence.
ACCESSION   AZ424155
VERSION     AZ424155.1  GI:10548168
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 100)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0203  row: C  column: 08
            Seq primer: CACACAGGAAACAGCTATGACC
            Class: plasmid ends
            High quality sequence stop: 100.
FEATURES             Location/Qualifiers
     source          1. .100
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC1M0203C08"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
BASE COUNT       32 a     13 c     22 g     33 t
ORIGIN

  Query Match            69.0%;  Score 13.8;  DB 17;  Length 100;
  Best Local Similarity  88.2%;  Pred. No. 5e+04;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 TCCTCTGGCTTCTGGTA 17
           |||||||||| | ||||
Db       7 TCCTCTGGCTACAGGTA 23
```

```
RESULT 10
AI430606/c
LOCUS       AI430606                 70 bp    mRNA    linear   EST 15-MAR-2000
DEFINITION  me08a02.y1 Soares mouse embryo NbME13.5 14.5 Mus musculus cDNA
            clone IMAGE:386858 5' similar to SW:Y195_HUMAN Q12767 HYPOTHETICAL
            PROTEIN KIAA0195. ;, mRNA sequence.
ACCESSION   AI430606
VERSION     AI430606.1  GI:4276442
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 70)
  AUTHORS   Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,
            Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person
            ,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter
            ,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R.,
            Waterston,R. and Wilson,R.
  TITLE     The WashU-NCI Mouse EST Project 1999
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Marra M/WashU-NCI Mouse EST Project 1999
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            This read is a RESEQUENCE of a previously sequenced mouse clone
            This read has been verified (found to hit its original self in the
            correct orientation)
            Trace considered overall poor quality
            Possible reversed clone: similarity on wrong strand
            MGI:238690
            Seq primer: -40RP from Gibco
            High quality sequence stop: 1
            POLYA=No.
FEATURES             Location/Qualifiers
     source          1. .70
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:386858"
                     /clone_lib="Soares mouse embryo NbME13.5 14.5"
                     /sex="unknown"
                     /tissue_type="embryo"
                     /dev_stage="13.5-14.5dpc total fetus"
                     /lab_host="DH10B"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                     was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCGGAAATTTTTTTTTTTTTTTTTTTTTT
                     T 3'], on equal amounts of mRNA from 2 13.5dpc and 2
                     14.5dpc embryos [total RNA provided by  Minoru Ko, Wayne
                     State Univ., from 2 ]; double-stranded cDNA was ligated to
                     Eco RI adaptors (Pharmacia), digested with Not I and
                     cloned into the Not I and Eco RI sites of the modified
                     pT7T3 vector. Library went through one round of
                     normalization, and was constructed by Bento Soares and
                     M.Fatima Bonaldo.  "
BASE COUNT       18 a     22 c     18 g     12 t
ORIGIN

  Query Match            68.0%;  Score 13.6;  DB 9;  Length 70;
  Best Local Similarity  80.0%;  Pred. No. 5.2e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 TCCTCTGGCTTCTGGTAGCG 20
           | ||   ||||||||||||
Db      30 TTCTGGCGCTTCTGGTAGCG 11
```

```
RESULT 11
AA894524/c
LOCUS       AA894524                 52 bp    mRNA    linear   EST 06-APR-1998
DEFINITION  of90c08.s1 NCI_CGAP_Li5 Homo sapiens cDNA clone IMAGE:1437614 3'
            similar to TR:Q84649 Q84649 GENOME, PARTIAL SEQUENCE. ;, mRNA
            sequence.
ACCESSION   AA894524
VERSION     AA894524.1  GI:3030925
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 52)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
            R. Emmert-Buck, M.D., Ph.D.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: Greg Lennon, Ph.D.
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html

            Trace considered overall poor quality
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .52
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1437614"
                     /clone_lib="NCI_CGAP_Li5"
                     /tissue_type="hepatic adenoma"
                     /lab_host="DH10B"
                     /note="Organ: liver; Vector: pCMV-SPORT4; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 0.8 kb."
BASE COUNT       26 a      4 c     22 g      0 t
ORIGIN

  Query Match             67.0%;  Score 13.4;  DB 9;  Length 52;
  Best Local Similarity   93.3%;  Pred. No. 5.6e+04;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 TCCTCTGGCTTCTGG 15
           ||||||||||||| |
Db      46 TCCTCTGGCTTCTCG 32


RESULT 12
AU258999
LOCUS       AU258999                 75 bp    mRNA    linear   EST 25-APR-2002
DEFINITION  AU258999 3'-directed mouse cDNA library Mus musculus cDNA clone
            BED0014316 3', mRNA sequence.
ACCESSION   AU258999
VERSION     AU258999.1  GI:20325104
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 75)
  AUTHORS   Kato,K. and Matoba,R.
  TITLE     Generation of expressed sequence tags from mouse brain
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Kikuya Kato
            Graduate School of Biological Sciences
            Nara Institute of Science and Technology
            8916-5 Takayama, Ikoma, Nara 630-0101, Japan
            Tel: 81-743-72-5581
            Fax: 81-743-72-5589
            Email: kkato@bs.aist-nara.ac.jp,
            URL:http://love2.aist-nara.ac.jp/BED/index.html.
FEATURES             Location/Qualifiers
     source          1. .75
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /clone="BED0014316"
                     /clone_lib="3'-directed mouse cDNA library"
                     /tissue_type="brain"
                     /note="Vector: pGEM-T-easy"
BASE COUNT       17 a      20 c      11 g      27 t
ORIGIN

  Query Match             67.0%;  Score 13.4;  DB 9;  Length 75;
  Best Local Similarity   93.3%;  Pred. No. 6.6e+04;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 TCCTCTGGCTTCTGG 15
           || ||||||||||||
Db      13 TCTTCTGGCTTCTGG 27


RESULT 13
AI120925/c
LOCUS       AI120925                 94 bp    mRNA    linear   EST 02-SEP-1998
DEFINITION  ub74d04.r1 Soares_mammary_gland_NMLMG Mus musculus cDNA clone
            IMAGE:1383463 5' similar to TR:O15554 O15554 INTERMEDIATE
            CONDUCTANCE CALCIUM-ACTIVATED POTASSIUM CHANNEL. ;, mRNA sequence.
ACCESSION   AI120925
VERSION     AI120925.1  GI:3521249
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 94)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:905931
            Trace considered overall poor quality
            Possible reversed clone: similarity on wrong strand
            Seq primer: -28m13 rev2 ET from Amersham
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .94
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:1383463"
                     /clone_lib="Soares_mammary_gland_NMLMG"
                     /sex="female (lactating)"
                     /tissue_type="mammary gland"
                     /lab_host="DH10B"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
```

polylinker; 1st strand cDNA was prepared from mammary gland tissue from a lactating female, and was then primed with a Not I - oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library is normalized. Library was constructed by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT 24 a 27 c 30 g 13 t
ORIGIN

Query Match 67.0%; Score 13.4; DB 9; Length 94;
Best Local Similarity 93.3%; Pred. No. 7.2e+04;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 5 CTGGCTTCTGGTAGC 19
|||| ||||||||||
Db 87 CTGGGTTCTGGTAGC 73

RESULT 14
BG550793/c
LOCUS BG550793 96 bp mRNA linear EST 28-NOV-2001
DEFINITION sad24a03.y1 Gm-c1074 Glycine max cDNA clone GENOME SYSTEMS CLONE ID: Gm-c1074-1157 5' similar to TR:Q41746 Q41746 CHLOROPHYLL A/B-BINDING APOPROTEIN CP26 PRECURSOR. ;, mRNA sequence.
ACCESSION BG550793
VERSION BG550793.1 GI:13562573
KEYWORDS EST.
SOURCE soybean.
ORGANISM Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
REFERENCE 1 (bases 1 to 96)
AUTHORS Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Coryell,V., Khanna ,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers ,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk ,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann ,R., Waterston,R. and Wilson,R.
TITLE Public Soybean EST Project
JOURNAL Unpublished (1999)
COMMENT Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Trace considered overall poor quality This clone is available through: ResGen, Invitrogen Corp. 2130 South Memorial Parkway Huntsville, AL 35801 For further information call: (800)-533-4363 or contact via email: ccu@resgen.com
High quality sequence stop: 1.
FEATURES Location/Qualifiers
source 1. .96
/organism="Glycine max"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-c1074-1157"
/clone_lib="Gm-c1074"
/tissue_type="seedlings induced for HR (hypersensitive response)"
/dev_stage="9-11 day old"
/lab_host="DH10B"
/note="Vector: pBluescript II SK+; Site_1: EcoRI; Site_2: XhoI; The cDNA library was constructed from mRNA isolated from 9-11 day old seedlings that were induced for HR (hypersensitive response) by vacuum infiltrating plant tissue with Pseudomonas syringae pv. glycinea carrying the avrB gene (Genetics 141:1597-1604). Plant tissue (expanded unifoliate leaves) was collected at 2, 4, 8, 12, 24, 36, and 53 hrs after inoculation and their mRNA pooled equally for cDNA construction. The library was prepared using the Stratagene pBluescript II SK(+) library construction kit. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dT) sequence with an XhoI restriction site. EcoRI adaptors were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA insert is protected from XhoI digestion via methylation during first strand synthesis. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into E.coli ElectroMax DH10B host cells. Plant care, inoculations, and library construction were performed by Steve Clough (Lila Vodkin lab, University of Illinois)."
BASE COUNT 25 a 23 c 27 g 21 t
ORIGIN

Query Match 67.0%; Score 13.4; DB 12; Length 96;
Best Local Similarity 93.3%; Pred. No. 7.3e+04;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TCCTCTGGCTTCTGG 15
||||||||||||| |
Db 87 TCCTCTGGCTTCTTG 73

RESULT 15
AZ820877/c
LOCUS AZ820877 99 bp DNA linear GSS 20-FEB-2001
DEFINITION 2M0093G21F Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC2M0093G21 F, DNA sequence.
ACCESSION AZ820877
VERSION AZ820877.1 GI:12990785
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 99)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D.,Weiss,R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0093 row: G column: 21
Seq primer: CGTTGTAAAACGACGGCCAGT
Class: plasmid ends
High quality sequence stop: 99.
FEATURES Location/Qualifiers
source 1. .99
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC2M0093G21"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a

0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptored DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptored mouse DNA was annealed to adaptored vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 34 a 15 c 13 g 37 t
ORIGIN

Query Match 67.0%; Score 13.4; DB 17; Length 99;
Best Local Similarity 93.3%; Pred. No. 7.4e+04;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy       6 TGGCTTCTGGTAGCG 20
           ||||||| ||||||
Db      22 TGGCTTCTTGTAGCG 8
```

Search completed: November 26, 2002, 04:08:58
Job time : 767.8 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 05:52:31 ; Search time 98.55 Seconds
(without alignments)
457.027 Million cell updates/sec

Title: US-09-296-264-17
Perfect score: 20
Sequence: 1 aggtttccttttccgatttc 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 2390332

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	20	100.0	20	21	AAZ31447	Human neuropilin m
c	2	14.8	74.0	29	17	AAT09344	Activator Ac PCR p
	3	14.8	74.0	50	19	AAV53101	Human IL4 promoter
	4	14.8	74.0	50	19	AAV53102	Human IL4 promoter
	5	14.2	71.0	50	20	AAZ20135	Human BRCA1 gene P
c	6	13.8	69.0	20	20	AAX61154	Human chromosome a
c	7	13.8	69.0	54	16	AAT21823	Human gene signatu
	8	13.8	69.0	89	21	AAC22201	Human secreted pro
c	9	13.8	69.0	96	24	ABK76764	Bacillus lichenifo
c	10	13.6	68.0	50	22	AAL29345	Human SNP oligonuc
c	11	13.6	68.0	93	16	AAT26341	Human gene signatu
c	12	13.4	67.0	85	21	AAC32255	Human secreted pro
c	13	13.4	67.0	89	22	AAL16884	Human breast cance
	14	13.4	67.0	99	24	ABL72545	Corn tassel-derive
	15	13.2	66.0	20	20	AAZ03703	PCR primer used to
c	16	13.2	66.0	20	20	AAZ01760	PCR primer used to
c	17	13.2	66.0	20	20	AAX56999	Ras gene modulatin
c	18	13.2	66.0	20	20	AAX21636	Human N-ras specif
c	19	13.2	66.0	20	21	AAA95874	Human N-ras antise
c	20	13.2	66.0	27	19	AAV96000	Solanidine glucosy
	21	13.2	66.0	30	22	AAF90557	Dissociation trans
	22	13.2	66.0	30	24	AAD31638	Arabidopsis thalia
	23	13.2	66.0	33	13	AAQ28043	Primer CASOL16. S
	24	13.2	66.0	38	16	AAQ88976	VEGF RNA nucleic a
	25	13.2	66.0	51	18	AAV76657	Staphylococcus aur
c	26	13.2	66.0	51	21	AAA76638	Human clone cg2530
c	27	13.2	66.0	51	21	AAA76639	Human clone cg2530
	28	13.2	66.0	54	14	AAQ46836	T4 gene 23 termina
	29	13.2	66.0	57	9	AAN81440	Sequence of yeast
	30	13.2	66.0	57	11	AAQ04915	Wild type signal s
	31	13.2	66.0	57	11	AAQ04916	Modified signal se
	32	13.2	66.0	57	11	AAQ04917	Modified signal se
	33	13.2	66.0	57	11	AAQ04918	Modified signal se
c	34	13.2	66.0	60	24	ABN45778	Human spliced tran
	35	13.2	66.0	63	11	AAQ04920	Modified signal se
	36	13.2	66.0	63	24	ABK32975	Yeast DNA sequence
	37	13.2	66.0	66	8	AAN71364	Sequence encoding
	38	13.2	66.0	66	11	AAQ04921	Modified signal se
	39	13.2	66.0	69	14	AAQ42522	Sequence encoding
	40	13.2	66.0	75	5	AAN40005	Yeast preinvertase
	41	13.2	66.0	75	5	AAN40006	Invertase signal s
	42	13.2	66.0	75	5	AAN40007	Invertase signal s
	43	13.2	66.0	75	5	AAN40008	Invertase signal s
	44	13.2	66.0	75	8	AAN71270	Sequence encoding
	45	13.2	66.0	75	12	AAQ11885	Invertase signal -

ALIGNMENTS

RESULT 1
AAZ31447
ID AAZ31447 standard; DNA; 20 BP.
XX
AC AAZ31447;
XX
DT 07-FEB-2000 (first entry)
XX
DE Human neuropilin mRNA specific antisense oligo GTI3618.
XX
KW Neuropilin; human; growth; metastasis; tumor; neovascularisation;
KW cancer; papilloma; diabetic retinopathy; antisense; ss.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9955855-A2.
XX
PD 04-NOV-1999.
XX
PF 23-APR-1999; 99WO-CA00324.
XX
PR 23-APR-1998; 98US-0082791.
XX
PA (GENE-) GENESENSE TECHNOLOGIES INC.
XX
PI Wright JA, Young AH, Lee YS;
XX
DR WPI; 2000-023357/02.
XX
PT Antisense oligonucleotides that inhibit neuropilin expression, useful
PT for treating cancer -

XX
PS Claim 4; Page 16; 57pp; English.
XX
CC Sequences AAZ31431-460 represent antisense oligonucleotides which
CC inhibit human neuropilin expression. The antisense oligonucleotides can
CC be used to inhibit the growth or metastasis of a mammalian tumor and
CC inhibit neovascularisation. The oligonucleotides may be used to treat
CC various forms of cancers or tumors, such as sarcomas, melanomas,
CC adenomas, carcinomas of solid tissue, hypoxic tumors, squamous cell
CC carcinomas of the mouth, throat, larynx and lung, genitourinary cancers
CC such as cervical and bladder cancer, hematopoietic cancers, colon cancer,
CC breast cancer, pancreatic cancer, renal cancer, brain cancer, skin
CC cancer, liver cancer, head and neck cancers, and nervous system cancers,
CC as well as benign lesions such as papillomas. The methods may be used to
CC treat neovascularisation disorders such as diabetic retinopathy, and
CC retinopathy of prematurity and age related macular degeneration.
XX
SQ Sequence 20 BP; 2 A; 5 C; 3 G; 10 T; 0 other;

Query Match 100.0%; Score 20; DB 21; Length 20;
Best Local Similarity 100.0%; Pred. No. 6;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 AGGTTTCCTTTTCCGATTTC 20
           ||||||||||||||||||||
Db       1 AGGTTTCCTTTTCCGATTTC 20
```

RESULT 2
AAT09344/c
ID AAT09344 standard; DNA; 29 BP.
XX
AC AAT09344;
XX
DT 28-MAR-1996 (first entry)
XX
DE Activator Ac PCR primer CC32.
XX
KW Tobacco mosaic virus resistance; TMV; N gene; Solanaceae;
KW crop improvement; transgenic plant; crop improvement;
KW Activator Ac; transposon; PCR; polymerase chain reaction;
KW primer; ss.
XX
OS Synthetic.
XX
PN WO9535024-A1.
XX
PD 28-DEC-1995.
XX
PF 16-JUN-1995; 95WO-US07754.
XX
PR 17-JUN-1994; 94US-0261663.
XX
PA (REGC) UNIV CALIFORNIA.
PA (USDA) US SEC OF AGRIC.
XX
PI Baker BJ, Whitham SA;
XX
DR WPI; 1996-058144/06.
XX
PT Plant virus resistance gene N sequences from tobacco - useful for
PT generating transgenic Solanaceous plants resistant to Tobacco Mosaic
PT Virus
XX
PS Example 3; Page 37; 98pp; English.
XX
CC A 419 bp product (Ac10-1) 5' to transposon Ac10 was amplified
CC using Ac specific primers CC28 and CC32 (AAT09343-44) using template
CC DNA from plant D111-95. 3' Flanking sequences were obtd. using
CC primers CC21 and CC30 (AAT09345-46). Sequences flanking Ac10 were
CC used to screen Nicotiana glutinosa cDNA and genomic DNA libraries
CC for clones coding for the N gene protein responsible for resistance
CC to tobacco mosaic virus.
XX
SQ Sequence 29 BP; 11 A; 6 C; 9 G; 3 T; 0 other;

Query Match 74.0%; Score 14.8; DB 17; Length 29;
Best Local Similarity 88.9%; Pred. No. 1.4e+03;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy       3 GTTTCCTTTTCCGATTTC 20
           |||||| ||||||| |||
Db      21 GTTTCCGTTTCCGAATTC 4
```

RESULT 3
AAV53101
ID AAV53101 standard; DNA; 50 BP.
XX
AC AAV53101;
XX
DT 12-NOV-1998 (first entry)
XX
DE Human IL4 promoter regions spanning probe (225-176).
XX
KW CP2 recognition element; IL4; promoter; asthma; therapeutic composition;
KW CP2 function affector; Th1/Th2 cell balance regulation; immune response;
KW immunological disease; allergic rhinitis, allergic conjunctivitis; CPRE;
KW dermatitis; urticaria; multiple sclerosis; arthritis; malignancy;
KW type I diabetes mellitus; parasitic infection; immunodeficient disorder;
KW T helper cell response; viral antigen; probe; ss.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9836641-A1.
XX
PD 27-AUG-1998.
XX
PF 19-FEB-1998; 98WO-US03049.
XX
PR 20-FEB-1997; 97US-0037972.
XX
PA (JOHN-) JOHNS HOPKINS SCHOOL MEDICINE.
PA (SCHE-) SCHEPENS EYE RES INST INC.
PA (SLOK) SLOAN KETTERING INST CANCER RES.
XX
PI Casolaro V, Ono SJ, Sheffery M, Swendeman SL;
XX
DR WPI; 1998-467194/40.
XX
PT Use of affector(s) of CP2 function - for modulating immune responses
PT for treating e.g. allergies, auto-immune disease, infections,
PT immunodeficiency disorders or malignancies
XX
PS Example 6; Page 23; 58pp; English.
XX
CC Sequences shown in AAV53101 to AAV53105 represent probes spanning the
CC human IL4 promoter regions which are used for binding experiments in the
CC course of the invention. The invention provides a therapeutic
CC composition comprising an affector of CP2 function in a carrier substance
CC and also a method of screening for such a CP2 function affector. The
CC method comprises providing first and second samples of components for an
CC assay for complex formation between CP2 and a CP2 recognition element
CC (CPRE) in the human IL4 promoter and causing the first sample of
CC components to react in the assay, where the extent of complex formation
CC between CP2 and a CPRE in the human IL4 promoter in the first assay
CC sample is determined. A candidate affector is added to the second sample
CC of components which is then caused to react in the assay, and the extent
CC of complex formation between CP2 and a CPRE in the human IL4 promoter in
CC the second assay sample is determined. The extent of complex formation
CC between the two assay samples is compared to determine the effect of the
CC candidate affector. The therapeutic composition comprising the affector
CC is used for the interruption or enhancement of CP2 activity and thus
CC regulation of Th1/Th2 cell balance, for therapeutic control of the immune
CC response and immunological disease in a variety of conditions including

CC allergic rhinitis, allergic conjunctivitis, asthma, dermatitis,
CC urticaria, multiple sclerosis, type I diabetes mellitus, arthritis and
CC parasitic infection. CP2 or dominant negative CP2 may also be useful in
CC the management of immunodeficient disorders or malignancies by
CC amplifying T helper cell responses to viral antigen.
XX
SQ Sequence 50 BP; 13 A; 5 C; 8 G; 24 T; 0 other;

Query Match 74.0%; Score 14.8; DB 19; Length 50;
Best Local Similarity 88.9%; Pred. No. 1.5e+03;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AGGTTTCCTTTTCCGATT 18
||||||| |||||| |||
Db 31 AGGTTTCATTTTCCTATT 48

RESULT 4
AAV53102
ID AAV53102 standard; DNA; 50 BP.
XX
AC AAV53102;
XX
DT 12-NOV-1998 (first entry)
XX
DE Human IL4 promoter regions spanning probe (195-146).
XX
KW CP2 recognition element; IL4; promoter; asthma; therapeutic composition;
KW CP2 function affector; Th1/Th2 cell balance regulation; immune response;
KW immunological disease; allergic rhinitis, allergic conjunctivitis; CPRE;
KW dermatitis; urticaria; multiple sclerosis; arthritis; malignancy;
KW type I diabetes mellitus; parasitic infection; immunodeficient disorder;
KW T helper cell response; viral antigen; probe; ss.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9836641-A1.
XX
PD 27-AUG-1998.
XX
PF 19-FEB-1998; 98WO-US03049.
XX
PR 20-FEB-1997; 97US-0037972.
XX
PA (JOHN-) JOHNS HOPKINS SCHOOL MEDICINE.
PA (SCHE-) SCHEPENS EYE RES INST INC.
PA (SLOK) SLOAN KETTERING INST CANCER RES.
XX
PI Casolaro V, Ono SJ, Sheffery M, Swendeman SL;
XX
DR WPI; 1998-467194/40.
XX
PT Use of affector(s) of CP2 function - for modulating immune responses
PT for treating e.g. allergies, auto-immune disease, infections,
PT immunodeficiency disorders or malignancies
XX
PS Example 6; Page 23; 58pp; English.
XX
CC Sequences shown in AAV53101 to AAV53105 represent probes spanning the
CC human IL4 promoter regions which are used for binding experiments in the
CC course of the invention. The invention provides a therapeutic
CC composition comprising an affector of CP2 function in a carrier substance
CC and also a method of screening for such a CP2 function affector. The
CC method comprises providing first and second samples of components for an
CC assay for complex formation between CP2 and a CP2 recognition element
CC (CPRE) in the human IL4 promoter and causing the first sample of
CC components to react in the assay, where the extent of complex formation
CC between CP2 and a CPRE in the human IL4 promoter in the first assay
CC sample is determined. A candidate affector is added to the second sample
CC of components which is then caused to react in the assay, and the extent
CC of complex formation between CP2 and a CPRE in the human IL4 promoter in
CC the second assay sample is determined. The extent of complex formation
CC between the two assay samples is compared to determine the effect of the
CC candidate affector. The therapeutic composition comprising the affector
CC is used for the interruption or enhancement of CP2 activity and thus
CC regulation of Th1/Th2 cell balance, for therapeutic control of the immune
CC response and immunological disease in a variety of conditions including
CC allergic rhinitis, allergic conjunctivitis, asthma, dermatitis,
CC urticaria, multiple sclerosis, type I diabetes mellitus, arthritis and
CC parasitic infection. CP2 or dominant negative CP2 may also be useful in
CC the management of immunodeficient disorders or malignancies by
CC amplifying T helper cell responses to viral antigen.
XX
SQ Sequence 50 BP; 10 A; 9 C; 8 G; 23 T; 0 other;

Query Match 74.0%; Score 14.8; DB 19; Length 50;
Best Local Similarity 88.9%; Pred. No. 1.5e+03;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AGGTTTCCTTTTCCGATT 18
||||||| |||||| |||
Db 1 AGGTTTCATTTTCCTATT 18

RESULT 5
AAZ20135
ID AAZ20135 standard; DNA; 50 BP.
XX
AC AAZ20135;
XX
DT 05-JAN-2000 (first entry)
XX
DE Human BRCA1 gene PCR primer.
XX
KW BRCA1; p53 protein; p21 gene; human; tumour suppressor;
KW transcriptional activator; breast cancer; cell proliferation;
KW apoptosis; diagnosis; anticancer; antitumour; drug screening;
KW PCR; primer; ss.
XX
OS Homo sapiens.
XX
PN WO9950280-A1.
XX
PD 07-OCT-1999.
XX
PF 31-MAR-1999; 99WO-US07150.
XX
PR 31-MAR-1998; 98US-0080146.
XX
PA (UYPE-) UNIV PENNSYLVANIA.
XX
PI El-Deiry WS, Weber BL;
XX
DR WPI; 1999-601319/51.
XX
PT Nucleic acid involved in BRCA-1-mediated control of transcriptional
PT regulation of tumour suppressor genes and related peptides, used to
PT screen for modulators for use as anticancer agents -
XX
PS Example 1; Page 31; 94pp; English.
XX
CC This primer is one of a pair (see also AAZ20134) used in the PCR
CC amplification of DNA encoding a fusion between the extreme
CC C-terminus of human tumour suppressor protein BRCA1 and the NLS
CC region (amino acids 499-510). The product was used in experiments
CC demonstrating that BRCA1 arrests the cell cycle by transactivating
CC the expression of the CDK-inhibitor p21. This newly discovered
CC pathway of BRCA1 action provides a novel target to which
CC pharmaceutical agents capable of affecting cell proliferation and
CC apoptosis can be isolated, and can then be used in the treatment
CC and control of cellular proliferation disorders, such as breast
CC cancer. Assays and compositions for identifying compounds that
CC enhance or repress cellular proliferation via this BRCA1-mediated
CC pathway are disclosed.
XX

```
SQ   Sequence 50 BP; 9 A; 7 C; 11 G; 23 T; 0 other;

  Query Match           71.0%;  Score 14.2;  DB 20;  Length 50;
  Best Local Similarity  84.2%;  Pred. No. 2.8e+03;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 AGGTTTCCTTTTCCGATTT 19
           |||| ||||||| || |||
Db      12 AGGTCTCCTTTTACGCTTT 30


RESULT 6
AAX61154/c
ID   AAX61154 standard; DNA; 20 BP.
XX
AC   AAX61154;
XX
DT   28-JUL-1999  (first entry)
XX
DE   Human chromosome alpha-satellite region.
XX
KW   Probe; human; chromosome 17 triple-helix forming oligonucleotide;
KW   genetic disorder; missing chromosome; aneuploidy; chromosome 21;
KW   infectious disease; diagnosis; alpha-satellite region; ss.
XX
OS   Homo sapiens.
XX
PN   WO9924622-A1.
XX
PD   20-MAY-1999.
XX
PF   10-NOV-1998;   98WO-US23765.
XX
PR   10-NOV-1997;   97US-0064997.
XX
PA   (UYPR-) UNIV PRINCETON.
XX
PI   Fresco JR,  Johnson MD;
XX
DR   WPI; 1999-327425/27.
XX
PT   Novel use of triple helix forming oligonucleotides, useful for in
PT   situ detection of double stranded target sequence
XX
PS   Claim 19; Page 12; 45pp; English.
XX
CC   This sequence represents a human chromosome alpha-satellite region.
CC   The invention relates to the use of a triple-helix forming
CC   oligonucleotide for in situ detection of a double-stranded target nucleic
CC   acid sequence. The method can be used to detect a genetic disorder
CC   e.g. to detect an extra or missing chromosome or fragment or aneuploidy,
CC   especially for detecting an extra or missing chromosome 17 or 21. The
CC   method can be also be used to screen for individuals at risk of
CC   developing a disease or for diagnosing an infectious disease. The use of
CC   triple helix forming oligonucleotides allows in situ detection of double
CC   stranded target sequence as opposed to prior art uses of developing
CC   potential anti-gene therapeutic agents or artificial restriction
CC   endonucleases.
XX
SQ   Sequence 20 BP; 11 A; 0 C; 7 G; 2 T; 0 other;

  Query Match           69.0%;  Score 13.8;  DB 20;  Length 20;
  Best Local Similarity  88.2%;  Pred. No. 4e+03;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       4 TTTCCTTTTCCGATTTC 20
           |||||||||||  ||||
Db      20 TTTCCTTTTCCATTTTC 4


RESULT 7
AAT21823/c
ID   AAT21823 standard; cDNA to mRNA; 54 BP.
XX
AC   AAT21823;
XX
DT   01-AUG-1996  (first entry)
XX
DE   Human gene signature HUMGS03313.
XX
KW   Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW   human; cloning; mapping; non-biased library; diagnosis; detection;
KW   cell typing; abnormal cell function; ss.
XX
OS   Homo sapiens.
XX
PN   WO9514772-A1.
XX
PD   01-JUN-1995.
XX
PF   11-NOV-1994;   94WO-JP01916.
XX
PR   12-NOV-1993;   93JP-0355504.
XX
PA   (MATS/) MATSUBARA K.
PA   (OKUB/) OKUBO K.
XX
PI   Matsubara K,  Okubo K;
XX
DR   WPI; 1995-206931/27.
XX
PT   Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT   for diagnosis of abnormal cell function, by preparing cDNA that
PT   reflects relative abundance of corresp. mRNA in specific human
PT   tissues
XX
PS   Claim 1; Page 974; 2245pp; Japanese.
XX
CC   A single-stranded DNA (or its complementary strand or the corresp.
CC   double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC   given in AAT19001-T26837 and which is able to hybridise to part of
CC   human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC   sequences were obtained from 3'-directed cDNA libraries prepared
CC   from various human tissues; synthesis of cDNA was initiated from the
CC   3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC   untranslated sequence is unique to a particular mRNA species, almost
CC   all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC   is constructed so as to reflect accurately the relative abundance of
CC   different mRNAs in the particular tissue from which it was derived.
CC   The appearance frequency of a given GS in a cDNA library can be
CC   determined (esp. using primers and probes derived from the GS
CC   sequences) as a means of diagnosing abnormal cell function or for
CC   recognising different cell types.
XX
SQ   Sequence 54 BP; 26 A; 4 C; 10 G; 13 T; 1 other;

  Query Match           69.0%;  Score 13.8;  DB 16;  Length 54;
  Best Local Similarity  88.2%;  Pred. No. 4.3e+03;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       4 TTTCCTTTTCCGATTTC 20
           |||| |||||| |||||
Db      54 TTTCGTTTTCCTATTTC 38


RESULT 8
AAC22201
ID   AAC22201 standard; cDNA; 89 BP.
XX
AC   AAC22201;
XX
DT   06-OCT-2000  (first entry)
XX
DE   Human secreted protein 5' EST, SEQ ID NO: 26276.
XX
KW   Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
```

KW gene therapy; chromosome mapping; ss.
XX
OS Homo sapiens.
XX
PN EP1033401-A2.
XX
PD 06-SEP-2000.
XX
PF 21-FEB-2000; 2000EP-0200610.
XX
PR 26-FEB-1999; 99US-0122487.
XX
PA (GEST) GENSET.
XX
PI Dumas Milne Edwards J, Duclert A, Giordano J;
XX
DR WPI; 2000-500381/45.
XX
PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX
PS Claim 1; SEQ ID 26276; 71pp + CD-ROM; English.
XX
CC The present sequence is one of a large number of 5' ESTs derived from
CC mRNAs encoding secreted proteins. No ORF has yet been conclusively
CC identified within the present sequence. The 5' ESTs were prepared from
CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
CC sequences usually correspond mainly to the 3' untranslated region (UTR)
CC of the mRNA because they are often obtained from oligo-dT primed cDNA
CC libraries. Such ESTs are not well suited for isolating cDNA sequences
CC derived from the 5' ends of mRNAs and even in those cases where longer
CC cDNA sequences have been obtained, the full 5' UTR is rarely included.
CC 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be
CC used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used
CC in diagnostic, forensic, gene therapy and chromosome mapping procedures.
CC They are used to obtain upstream regulatory sequences and to design
CC expression and secretion vectors.
XX
SQ Sequence 89 BP; 24 A; 16 C; 12 G; 37 T; 0 other;

Query Match 69.0%; Score 13.8; DB 21; Length 89;
Best Local Similarity 88.2%; Pred. No. 4.4e+03;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 TTTCCTTTTCCGATTTC 20
||||||||||||| || |
Db 4 TTTCCTTTTCCGGTTGC 20

RESULT 9
ABK76764/c
ID ABK76764 standard; DNA; 96 BP.
XX
AC ABK76764;
XX
DT 13-AUG-2002 (first entry)
XX
DE Bacillus licheniformis genomic sequence tag (GST) #4055.
XX
KW Differential gene expression; genomic sequenced tag; GST;
KW altered culture condition; environmental stress;
KW physiological provocation; ds.
XX
OS Bacillus licheniformis.
XX
PN WO200229113-A2.
XX
PD 11-APR-2002.
XX
PF 05-OCT-2001; 2001WO-US31437.
XX
PR 06-OCT-2000; 2000US-0680598.
PR 27-MAR-2001; 2001US-279526P.
XX
PA (NOVO) NOVOZYMES BIOTECH INC.
PA (NOVO) NOVOZYMES AS.
XX
PI Berka R, Clausen IG;
XX
DR WPI; 2002-416684/44.
XX
PT Monitoring differential expression of several genes in first Bacillus
PT cell relative to expression of same genes in one or more second
PT Bacillus cells, by using substrate containing Bacillus genomic
PT sequenced tag array -
XX
PS Claim 4; SEQ ID NO 4055; 200pp; English.
XX
CC The invention describes a method of monitoring differential expression of
CC genes in a first Bacillus cell relative to expression of the genes in
CC other Bacillus cells, comprising hybridising labelled nucleic acid probes
CC isolated from Bacillus cells to a substrate containing array of Bacillus
CC genomic sequenced tags (GST), examining the array, and determining
CC relative gene expression by an observed hybridisation reporter signal of
CC a spot in the array. The method is useful for measuring the expression of
CC genes in a first Bacillus cell relative to expression of the same genes
CC in one or more second Bacillus cells. The method is useful for monitoring
CC global expression of several genes from a Bacillus cell, discovering new
CC genes, identifying possible functions of unknown open reading frames and
CC monitoring gene copy number variation and stability. Monitoring changes
CC in expression of genes may be used to provide a representation of the way
CC in which Bacillus cells adapt to changes in culture conditions,
CC environmental stress or other physiological provocation. Extensive
CC follow-up characterisation is unnecessary, when one spot on an array
CC equals one gene or one open reading frame, since sequence information is
CC available. This sequence represents a genomic sequence tag (GST) used in
CC the method of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 96 BP; 32 A; 25 C; 18 G; 21 T; 0 other;

Query Match 69.0%; Score 13.8; DB 24; Length 96;
Best Local Similarity 88.2%; Pred. No. 4.4e+03;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 TTTCCTTTTCCGATTTC 20
||||||||| | |||||
Db 39 TTTCCTTTTTCAATTTC 23

RESULT 10
AAL29345/c
ID AAL29345 standard; DNA; 50 BP.
XX
AC AAL29345;
XX
DT 24-JAN-2002 (first entry)
XX
DE Human SNP oligonucleotide #2553.
XX
KW Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
KW neuroprotective; antimicrobial; gene therapy; vaccine; amylase; cancer;
KW amyloid protein; angiopoietin; apoptosis related protein; cadherin;
KW cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
KW complement related protein; cytochrome; kinesin; cytokine; interferon;
KW interleukin; G-protein coupled receptor; thioesterase; inflammation;
KW multifactorial disease; autoimmune disease; infection;
KW nervous system disease; ss.
XX
OS Homo sapiens.
XX
PN WO200147944-A2.

```
XX
PD   05-JUL-2001.
XX
PF   28-DEC-2000; 2000WO-US35498.
XX
PR   28-DEC-1999;   99US-0173419.
PR   27-DEC-2000; 2000US-0173419.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Shimkets RA,  Leach M;
XX
DR   WPI; 2001-465210/50.
XX
PT   Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
PT   oncogenes and histones, useful for diagnosing and treating, e.g.
PT   cancer, autoimmune diseases and infections -
XX
PS   Claim 1; Page 2115; 4143pp; English.
XX
CC   The present invention relates to oligonucleotides encoding polymorphic
CC   variants of proteins related to amylases, amyloid proteins, angiopoietin,
CC   apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,
CC   histones, kinases, colony stimulating factors, complement related
CC   proteins, cytochromes, kinesins, cytokines, interferons, interleukins,
CC   G-protein coupled receptors and thioesterases. The present sequence is
CC   one such oligonucleotide. The oligonucleotides and the peptides encoded
CC   by them may be used in the prevention, diagnosis and treatment of
CC   diseases associated with inappropriate expression of the proteins listed
CC   above. Disorders that may be prevented, diagnosed and/or treated include
CC   multifactorial diseases with a genetic component, such as autoimmune
CC   diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,
CC   systemic lupus erythromatosus and Grave's disease), inflammation, cancer
CC   (e.g. cancers of the bladder, brain, breast, colon and kidney,
CC   leukaemia), diseases of the nervous system and an infection of pathogenic
CC   organisms.
XX
SQ   Sequence 50 BP; 25 A; 6 C; 8 G; 11 T; 0 other;

  Query Match             68.0%;  Score 13.6;  DB 22;  Length 50;
  Best Local Similarity   80.0%;  Pred. No. 5.2e+03;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 AGGTTTCCTTTTCCGATTTC 20
           ||  || ||||||| |||||
Db      45 AGCCTTTCTTTTCCTATTTC 26


RESULT 11
AAT26341/c
ID   AAT26341 standard; cDNA to mRNA; 93 BP.
XX
AC   AAT26341;
XX
DT   16-OCT-1996  (first entry)
XX
DE   Human gene signature HUMGS08581.
XX
KW   Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW   human; cloning; mapping; non-biased library; diagnosis; detection;
KW   cell typing; abnormal cell function; ss.
XX
OS   Homo sapiens.
XX
PN   WO9514772-A1.
XX
PD   01-JUN-1995.
XX
PF   11-NOV-1994;   94WO-JP01916.
XX
PR   12-NOV-1993;   93JP-0355504.
XX
PA   (MATS/) MATSUBARA K.
PA   (OKUB/) OKUBO K.
XX
PI   Matsubara K,  Okubo K;
XX
DR   WPI; 1995-206931/27.
XX
PT   Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT   for diagnosis of abnormal cell function, by preparing cDNA that
PT   reflects relative abundance of corresp. mRNA in specific human
PT   tissues
XX
PS   Claim 1; Page 2061; 2245pp; Japanese.
XX
CC   A single-stranded DNA (or its complementary strand or the corresp.
CC   double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC   given in AAT19001-T26837 and which is able to hybridise to part of
CC   human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC   sequences were obtained from 3'-directed cDNA libraries prepared
CC   from various human tissues; synthesis of cDNA was initiated from the
CC   3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC   untranslated sequence is unique to a particular mRNA species, almost
CC   all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC   is constructed so as to reflect accurately the relative abundance of
CC   different mRNAs in the particular tissue from which it was derived.
CC   The appearance frequency of a given GS in a cDNA library can be
CC   determined (esp. using primers and probes derived from the GS
CC   sequences) as a means of diagnosing abnormal cell function or for
CC   recognising different cell types.
XX
SQ   Sequence 93 BP; 34 A; 14 C; 15 G; 30 T; 0 other;

  Query Match             68.0%;  Score 13.6;  DB 16;  Length 93;
  Best Local Similarity   80.0%;  Pred. No. 5.5e+03;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 AGGTTTCCTTTTCCGATTTC 20
           || |||||||||   |||||
Db      83 AGATTTCCTTTTATTATTTC 64


RESULT 12
AAC32255/c
ID   AAC32255 standard; cDNA; 85 BP.
XX
AC   AAC32255;
XX
DT   06-OCT-2000  (first entry)
XX
DE   Human secreted protein 5' EST, SEQ ID NO: 36330.
XX
KW   Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW   gene therapy; chromosome mapping; ss.
XX
OS   Homo sapiens.
XX
PN   EP1033401-A2.
XX
PD   06-SEP-2000.
XX
PF   21-FEB-2000; 2000EP-0200610.
XX
PR   26-FEB-1999;   99US-0122487.
XX
PA   (GEST ) GENSET.
XX
PI   Dumas Milne Edwards J,  Duclert A,  Giordano J;
XX
DR   WPI; 2000-500381/45.
XX
PT   New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT   obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT   diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX
```

PS Claim 1; SEQ ID 36330; 71pp + CD-ROM; English.
XX
CC The present sequence is one of a large number of 5' ESTs derived from
CC mRNAs encoding secreted proteins. No ORF has yet been conclusively
CC identified within the present sequence. The 5' ESTs were prepared from
CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
CC sequences usually correspond mainly to the 3' untranslated region (UTR)
CC of the mRNA because they are often obtained from oligo-dT primed cDNA
CC libraries. Such ESTs are not well suited for isolating cDNA sequences
CC derived from the 5' ends of mRNAs and even in those cases where longer
CC cDNA sequences have been obtained, the full 5' UTR is rarely included.
CC 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be
CC used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used
CC in diagnostic, forensic, gene therapy and chromosome mapping procedures.
CC They are used to obtain upstream regulatory sequences and to design
CC expression and secretion vectors.
XX
SQ Sequence 85 BP; 38 A; 6 C; 37 G; 3 T; 1 other;

Query Match 67.0%; Score 13.4; DB 21; Length 85;
Best Local Similarity 82.4%; Pred. No. 6.7e+03;
Matches 14; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 4 TTTCCTTTTCCGATTTC 20
|||| ||:||||||||
Db 64 TTTCTCTTKCCGATTTC 48

RESULT 13
AAL16884/c
ID AAL16884 standard; cDNA; 89 BP.
XX
AC AAL16884;
XX
DT 07-DEC-2001 (first entry)
XX
DE Human breast cancer expressed polynucleotide 9341.
XX
KW Human; breast cancer; cell marker; cytostatic; ss.
XX
OS Homo sapiens.
XX
PN WO200151628-A2.
XX
PD 19-JUL-2001.
XX
PF 10-JAN-2001; 2001WO-US00798.
XX
PR 14-JAN-2000; 2000US-0176077.
PR 14-MAR-2000; 2000US-0189167.
PR 24-MAR-2000; 2000US-0192099.
PR 29-MAR-2000; 2000US-0193480.
PR 15-MAY-2000; 2000US-0205230.
PR 09-JUN-2000; 2000US-0211315.
PR 25-JUL-2000; 2000US-0220534.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Lillie J, Xu Y, Wang Y, Steinmann K;
XX
DR WPI; 2001-451856/48.
XX
PT New peptide useful as a marker for the diagnosis of breast cancer -
XX
PS Claim 1; Page 1676; 3695pp; English.
XX
CC The invention relates to human breast cancer expressed polynucleotides
CC (AAL07544-AAL26789) and methods of assessing whether a patient is
CC afflicted with breast cancer by examining the correlation between the
CC expression of certain markers and the cancerous state of breast cells.
CC The polynucleotides and encoded polypeptides are potential markers for
CC detecting, diagnosing, monitoring, characterising treating and
CC potentially preventing breast cancer. The polynucleotides and encoded
CC polypeptides are also useful for isolating compounds with cytostatic
CC activity.
XX
SQ Sequence 89 BP; 37 A; 18 C; 15 G; 19 T; 0 other;

Query Match 67.0%; Score 13.4; DB 22; Length 89;
Best Local Similarity 93.3%; Pred. No. 6.7e+03;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 GTTTCCTTTCCGAT 17
||||||||||| |||
Db 48 GTTTCCTTTTCAGAT 34

RESULT 14
ABL72545
ID ABL72545 standard; cDNA; 99 BP.
XX
AC ABL72545;
XX
DT 14-MAY-2002 (first entry)
XX
DE Corn tassel-derived polynucleotide (cdps) SEQ ID NO:1919.
XX
KW Corn; corn tassel-derived polynucleotide; cdps; hybrid breeding; CDPs;
KW inheritance; characteristic; growth; development; disease resistance;
KW environmental adaptability; quality; yield; molecular marker;
KW multigene trait; plant breeding; corn tassel; gene; ss.
XX
OS Zea mays.
XX
PN US2001051335-A1.
XX
PD 13-DEC-2001.
XX
PF 16-APR-1999; 99US-0294093.
XX
PR 21-APR-1998; 98US-082567P.
XX
PA (LALG/) LALGUDI R V.
PA (ITOL/) ITO L Y.
PA (SHER/) SHERMAN B K.
XX
PI Lalgudi RV, Ito LY, Sherman BK;
XX
DR WPI; 2002-163647/21.
XX
PT Novel purified corn tassel-derived polynucleotide useful for
PT determining altered gene expression, to recover regulatory elements and
PT to follow inheritance of desirable characteristics through hybrid
PT breeding programs -
XX
PS Claim 1; SEQ ID 1919; 201pp; English.
XX
CC The present sequence describes a purified corn tassel-derived
CC polynucleotide sequence (cdps) comprising a nucleic acid sequence
CC selected from those given in ABL70627 to ABL76833. The cdps sequences
CC encode corn tassel-derived polypeptides (CDPs). The cdps sequences (I)
CC can be used for determining altered gene expression, to recover
CC regulatory elements and to follow inheritance of desirable
CC characteristics through hybrid breeding programs. (I) are also useful
CC in the evaluation, and alteration of desired characteristics associated
CC with growth and development, disease resistance, environmental
CC adaptability, quality and yield, and as molecular markers for studying
CC inheritance of multigene traits in a plant breeding program. (I) can be
CC used to produce a tassel-specific profile of gene transcription, a
CC transcript image, to clone regulatory elements for use in transformation
CC vectors, to express a polypeptide, to identify, isolate or extend
CC identical or related corn tassel nucleic acid sequences from DNA
CC libraries, in nucleic acid hybridisation or amplification technologies,
CC as query sequences to determine homology of known sequences, as probe
CC for use in Southern or Northern hybridisation, and to identify the
CC presence of and/or to determine the degree of similarity between two

```
CC   (or more) nucleic acid sequences.
XX
SQ   Sequence 99 BP; 25 A; 25 C; 25 G; 17 T; 7 other;

  Query Match             67.0%;  Score 13.4;  DB 24;  Length 99;
  Best Local Similarity   93.3%;  Pred. No. 6.8e+03;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       6 TCCTTTTCCGATTTC 20
           ||||||||| |||||
Db       2 TCCTTTTCCAATTTC 16


RESULT 15
AAZ03703
ID   AAZ03703 standard; DNA; 20 BP.
XX
AC   AAZ03703;
XX
DT   07-OCT-1999  (first entry)
XX
DE   PCR primer used to amplify an ORF of Chlamydia trachomatis.
XX
KW   Vaccine; eye disease; conventional trachoma; nonendemic trachoma;
KW   paratrachoma; inclusion conjunctivitis; genital disease; perihepatitis;
KW   nongonococcal uretritis; epidymitis; cervicitis; salpingitis; PCR primer;
KW   bartholinitis; pneumopathy; venereal lymphogranulomatosis; ss.
XX
OS   Synthetic.
OS   Chlamydia trachomatis.
XX
PN   WO9928475-A2.
XX
PD   10-JUN-1999.
XX
PF   27-NOV-1998;   98WO-IB01939.
XX
PR   04-NOV-1998;   98US-0107077.
PR   28-NOV-1997;   97FR-0015041.
PR   17-DEC-1997;   97FR-0016034.
XX
PA   (GEST ) GENSET.
XX
PI   Griffais R;
XX
DR   WPI; 1999-371125/31.
XX
PT   Genome sequence of Chlamydia trachomatis
XX
PS   Disclosure; Page 1628; 1755pp; English.
XX
CC   PCR primers AAZ01426-Z06209 were used to amplify open reading frames
CC   (ORFs) of the genome of Chlamydia trachomatis (see AAZ01425). These ORFs
CC   encode polypeptides (see AAY36754-Y37949) which can be used as vaccines
CC   against Chlamydia trachomatis. Antisense and ribozyme sequences
CC   can also be used to control growth of the microorganism. Chlamydia
CC   trachomatis is responsible for a large number of diseases, e.g. eye
CC   diseases such as conventional trachoma, nonendemic trachoma,
CC   paratrachoma, and inclusion conjunctivitis; genital diseases such as
CC   nongonococcal uretritis, epidymitis, cervicitis, salpingitis,
CC   perihepatitis, bartholinitis; pneumopathy in breast feeding infants;
CC   and venereal lymphogranulomatosis. The polypeptides of the
CC   invention may be of use in treating these diseases.
XX
SQ   Sequence 20 BP; 4 A; 6 C; 4 G; 6 T; 0 other;

  Query Match             66.0%;  Score 13.2;  DB 20;  Length 20;
  Best Local Similarity   83.3%;  Pred. No. 7.5e+03;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 AGGTTTCCTTTTCCGATT 18
           |||  |||| ||||||||
Db       2 AGGAATCCTCTTCCGATT 19
```

Search completed: November 23, 2002, 06:29:22
Job time : 102.6 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 05:53:51 ; Search time 21.55 Seconds
(without alignments)
284.619 Million cell updates/sec

Title: US-09-296-264-17
Perfect score: 20
Sequence: 1 aggtttccttttccgatttc 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 687286

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	13.8	69.0	20	4	US-09-531-000-20	Sequence 20, Appl
	2	13.4	67.0	47	4	US-09-641-638-1162	Sequence 1162, Ap
c	3	13.2	66.0	20	2	US-08-889-296A-36	Sequence 36, Appl
c	4	13.2	66.0	20	3	US-08-961-469A-43	Sequence 43, Appl
c	5	13.2	66.0	20	3	US-09-128-494-36	Sequence 36, Appl
	6	13.2	66.0	25	4	US-09-102-528-14	Sequence 14, Appl
	7	13.2	66.0	34	1	US-08-104-072B-32	Sequence 32, Appl
	8	13.2	66.0	57	6	5221624-13	Patent No. 5221624
	9	13.2	66.0	57	6	5221624-15	Patent No. 5221624
	10	13.2	66.0	57	6	5221624-17	Patent No. 5221624
	11	13.2	66.0	57	6	5221624-19	Patent No. 5221624
	12	13.2	66.0	63	4	US-09-603-663-74	Sequence 74, Appl
	13	13.2	66.0	63	4	US-09-603-658-74	Sequence 74, Appl
	14	13.2	66.0	63	4	US-09-602-373A-74	Sequence 74, Appl
	15	13.2	66.0	63	6	5221624-23	Patent No. 5221624
	16	13.2	66.0	66	6	5221624-25	Patent No. 5221624
	17	13.2	66.0	69	1	US-08-244-492A-8	Sequence 8, Appli
	18	13.2	66.0	69	1	US-08-709-915-10	Sequence 10, Appl
	19	13.2	66.0	75	5	PCT-US94-09653A-32	Sequence 32, Appl
	20	13.2	66.0	77	1	US-08-447-169A-33	Sequence 33, Appl
	21	13.2	66.0	77	2	US-08-233-012C-33	Sequence 33, Appl
	22	13.2	66.0	84	2	US-08-369-829A-15	Sequence 15, Appl
	23	13.2	66.0	84	2	US-08-586-676E-20	Sequence 20, Appl
	24	13.2	66.0	84	5	PCT-US94-09653A-33	Sequence 33, Appl
	25	13.2	66.0	92	3	US-08-463-903-61	Sequence 61, Appl
	26	13.2	66.0	92	4	US-07-935-695-61	Sequence 61, Appl
c	27	12.8	64.0	26	2	US-08-690-734A-55	Sequence 55, Appl
c	28	12.8	64.0	26	3	US-08-742-185-55	Sequence 55, Appl
	29	12.8	64.0	29	1	US-08-105-761-10	Sequence 10, Appl
	30	12.8	64.0	29	5	PCT-US92-11076-10	Sequence 10, Appl
	31	12.8	64.0	74	3	US-08-910-632-30	Sequence 30, Appl
	32	12.8	64.0	74	3	US-08-805-631A-30	Sequence 30, Appl
	33	12.8	64.0	74	4	US-09-569-344-30	Sequence 30, Appl
	34	12.6	63.0	22	1	US-08-222-177A-153	Sequence 153, App
	35	12.6	63.0	24	4	US-09-434-774-3	Sequence 3, Appli
c	36	12.6	63.0	30	1	US-08-012-543-26	Sequence 26, Appl
c	37	12.6	63.0	30	2	US-08-468-558-24	Sequence 24, Appl
c	38	12.6	63.0	30	4	US-08-676-444-24	Sequence 24, Appl
c	39	12.6	63.0	31	2	US-09-018-576-8	Sequence 8, Appli
	40	12.6	63.0	31	2	US-09-018-576-9	Sequence 9, Appli
c	41	12.6	63.0	31	3	US-09-248-137-8	Sequence 8, Appli
	42	12.6	63.0	31	3	US-09-248-137-9	Sequence 9, Appli
c	43	12.6	63.0	57	2	US-08-811-492-119	Sequence 119, App
c	44	12.4	62.0	31	3	US-08-976-255-34	Sequence 34, Appl
c	45	12.2	61.0	20	1	US-08-671-947-35	Sequence 35, Appl

ALIGNMENTS

RESULT 1
US-09-531-000-20/c
; Sequence 20, Application US/09531000
; Patent No. 6461810
; GENERAL INFORMATION:
; APPLICANT: JOHNSON, Marion D.
; APPLICANT: FRESCO, Jacques R.
; TITLE OF INVENTION: TRIPLEX IN-SITU HYBRIDIZATION
; FILE REFERENCE: 2448-103
; CURRENT APPLICATION NUMBER: US/09/531,000
; CURRENT FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: PCT/US98/23765
; PRIOR FILING DATE: 1998-11-10
; PRIOR APPLICATION NUMBER: 60/064,997
; PRIOR FILING DATE: 1997-11-10
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 20
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Target
; OTHER INFORMATION: sequences
US-09-531-000-20

Query Match 69.0%; Score 13.8; DB 4; Length 20;
Best Local Similarity 88.2%; Pred. No. 3.6e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy       4 TTTCCTTTTCCGATTTC 20
           ||||||||||| ||||
Db      20 TTTCCTTTTCCATTTTC 4
```

RESULT 2
US-09-641-638-1162
; Sequence 1162, Application US/09641638
; Patent No. 6432648
; GENERAL INFORMATION:
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Chumakov, Ilya
; APPLICANT: Cohen, Annick
; TITLE OF INVENTION: BIALLELIC MARKERS DERIVED FROM GENOMIC REGIONS CARRYING
; TITLE OF INVENTION: GENES INVOLVED IN ARACHIDONIC ACID METABOLISM
; FILE REFERENCE: GENSET.051CP1
; CURRENT APPLICATION NUMBER: US/09/641,638
; CURRENT FILING DATE: 2000-08-16
; PRIOR APPLICATION NUMBER: US 09/502,330

```
;  PRIOR FILING DATE:  2000-02-11
;  PRIOR APPLICATION NUMBER:  US 60/133,200
;  PRIOR FILING DATE:  1999-05-07
;  PRIOR APPLICATION NUMBER:  US 09/275,267
;  PRIOR FILING DATE:  1999-03-23
;  PRIOR APPLICATION NUMBER:  US 60/119,917
;  PRIOR FILING DATE:  1999-02-12
;  NUMBER OF SEQ ID NOS:  1304
;  SOFTWARE:  Patent.pm
; SEQ ID NO 1162
;   LENGTH: 47
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 24
;   OTHER INFORMATION: 10-7-383 : polymorphic base C or T
US-09-641-638-1162

  Query Match             67.0%;  Score 13.4;  DB 4;  Length 47;
  Best Local Similarity   82.4%;  Pred. No. 6.1e+02;
  Matches   14;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       4 TTTCCTTTTCCGATTTC 20
           :| ||||||| ||||||
Db      24 YTACCTTTTCAGATTTC 40


RESULT 3
US-08-889-296A-36/c
; Sequence 36, Application US/08889296A
; Patent No. 5872242
;  GENERAL INFORMATION:
;    APPLICANT:  Monia, B.P., Cowsert, L.M. and Manoharan, M.
;    TITLE OF INVENTION:  Antisense Oligonucleotide
;    TITLE OF INVENTION:  Inhibition of ras
;    NUMBER OF SEQUENCES:  55
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Jane Massey Licata
;      STREET:  210 Lake Drive East, Suite 201
;      CITY:  Cherry Hill
;      STATE:  NJ
;      COUNTRY:  USA
;      ZIP:  08002
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  DISKETTE, 3.5 INCH, 1.44 Mb STORAGE
;      COMPUTER:  IBM PS/2
;      OPERATING SYSTEM:  PC-DOS
;      SOFTWARE:  WORDPERFECT 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/889,296A
;      FILING DATE:  herewith
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/411,734
;      FILING DATE:  April 3, 1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US93/09346
;      FILING DATE:  October 1, 1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  958,134
;      FILING DATE:  October 5, 1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/007,996
;      FILING DATE:  January 21, 1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jane Massey Licata
;      REGISTRATION NUMBER:  32,257
;      REFERENCE/DOCKET NUMBER:  ISPH-0213
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (609) 779-2400
;      TELEFAX:  (609) 779-8488
;  INFORMATION FOR SEQ ID NO:  36:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20
;      TYPE:  Nucleic Acid
;      STRANDEDNESS:  Single
;      TOPOLOGY:  Linear
;    ANTI-SENSE:  Yes
US-08-889-296A-36

  Query Match             66.0%;  Score 13.2;  DB 2;  Length 20;
  Best Local Similarity   83.3%;  Pred. No. 6.9e+02;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 GGTTTCCTTTTCCGATTT 19
           |||||| ||| | |||||
Db      19 GGTTTCTTTTACTGATTT 2


RESULT 4
US-08-961-469A-43/c
; Sequence 43, Application US/08961469A
; Patent No. 6083923
;  GENERAL INFORMATION:
;    APPLICANT:  Greg Hardee, Richard Geary, Arthur Levin,
;    APPLICANT:  Mike Templin, Randy Howard, Rahul Mehta
;    TITLE OF INVENTION:  LIPOSOMAL OLIGONUCLEOTIDE COMPOSITIONS           FOR MO
;    NUMBER OF SEQUENCES:  61
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Jane Massey Licata, Esq.
;      STREET:  66 E. Main Street
;      CITY:  Marlton
;      STATE:  NJ
;      COUNTRY:  USA
;      ZIP:  08053
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  DISKETTE, 3.5 INCH, 1.44 Mb STORAGE
;      COMPUTER:  PENTIUM
;      OPERATING SYSTEM:  WINDOWS 95
;      SOFTWARE:  WORDPERFECT 6.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/961,469A
;      FILING DATE:  October 31, 1997
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jane Massey Licata
;      REGISTRATION NUMBER:  32,257
;      REFERENCE/DOCKET NUMBER:  ISPH-0219
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  609-779-2400
;      TELEFAX:  609-810-1454
;  INFORMATION FOR SEQ ID NO:  43:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20
;      TYPE:  Nucleic Acid
;      STRANDEDNESS:  Single
;      TOPOLOGY:  Linear
;    ANTI-SENSE:  Yes
US-08-961-469A-43

  Query Match             66.0%;  Score 13.2;  DB 3;  Length 20;
  Best Local Similarity   83.3%;  Pred. No. 6.9e+02;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 GGTTTCCTTTTCCGATTT 19
           |||||| ||| | |||||
Db      19 GGTTTCTTTTACTGATTT 2


RESULT 5
US-09-128-494-36/c
; Sequence 36, Application US/09128494
```

```
; Patent No. 6117848
;  GENERAL INFORMATION:
;    APPLICANT:  Monia, B.P., Cowsert, L.M. and Manoharan, M.
;    TITLE OF INVENTION:  Antisense Oligonucleotide
;    TITLE OF INVENTION:  Inhibition of ras
;    NUMBER OF SEQUENCES:  55
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Jane Massey Licata
;      STREET:  210 Lake Drive East, Suite 201
;      CITY:  Cherry Hill
;      STATE:  NJ
;      COUNTRY:  USA
;      ZIP:  08002
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  DISKETTE, 3.5 INCH, 1.44 Mb STORAGE
;      COMPUTER:  IBM PS/2
;      OPERATING SYSTEM:  PC-DOS
;      SOFTWARE:  WORDPERFECT 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/128,494
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/889,296
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/411,734
;      FILING DATE:  April 3, 1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US93/09346
;      FILING DATE:  October 1, 1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  958,134
;      FILING DATE:  October 5, 1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/007,996
;      FILING DATE:  January 21, 1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jane Massey Licata
;      REGISTRATION NUMBER:  32,257
;      REFERENCE/DOCKET NUMBER:  ISPH-0213
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (609) 779-2400
;      TELEFAX:  (609) 779-8488
;  INFORMATION FOR SEQ ID NO:  36:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20
;      TYPE:  Nucleic Acid
;      STRANDEDNESS:  Single
;      TOPOLOGY:  Linear
;    ANTI-SENSE:  Yes
US-09-128-494-36

  Query Match             66.0%;  Score 13.2;  DB 3;  Length 20;
  Best Local Similarity   83.3%;  Pred. No. 6.9e+02;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 GGTTTCCTTTTCCGATTT 19
           |||||| ||| | |||||
Db      19 GGTTTCTTTTACTGATTT 2


RESULT 6
US-09-102-528-14
; Sequence 14, Application US/09102528
; Patent No. 6207883
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  DNA Sequences
;    NUMBER OF SEQUENCES:  32
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30 (EPO)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/102,528
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO PCT/GB96/03191
;      FILING DATE:
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  25 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-09-102-528-14

  Query Match             66.0%;  Score 13.2;  DB 4;  Length 25;
  Best Local Similarity   83.3%;  Pred. No. 7.1e+02;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       3 GTTTCCTTTTCCGATTTC 20
           |||||| |||||| || |
Db       2 GTTTCCGTTTCCGTTTAC 19


RESULT 7
US-08-104-072B-32
; Sequence 32, Application US/08104072B
; Patent No. 5639948
;  GENERAL INFORMATION:
;    APPLICANT:  Michiels, Frank
;    APPLICANT:  Morioka, Sinji
;    APPLICANT:  Scheirlinck, Trees
;    APPLICANT:  Komari, Tosihiko
;    TITLE OF INVENTION:  Stamen-specific Promoters from Rice
;    NUMBER OF SEQUENCES:  38
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Merchant & Gould
;      STREET:  3100 No. 5639948west Center
;      CITY:  Minneapolis
;      STATE:  MN
;      COUNTRY:  USA
;      ZIP:  55402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/104,072B
;      FILING DATE:  05-AUG-1993
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO 9200272
;      FILING DATE:  06-FEB-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  EP 91403352.7
;      FILING DATE:  10-DEC-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  EP 91402590.3
;      FILING DATE:  27-SEP-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  EP 91400318.1
;      FILING DATE:  08-FEB-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Kowalchyk, Katherine M.
;      REGISTRATION NUMBER:  36,848
;      REFERENCE/DOCKET NUMBER:  8076.93USWO
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  612-332-5300
;      TELEFAX:  612-332-9081
;  INFORMATION FOR SEQ ID NO:  32:
```

```
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  34 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-104-072B-32

  Query Match             66.0%;  Score 13.2;  DB 1;  Length 34;
  Best Local Similarity   83.3%;  Pred. No. 7.3e+02;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 GGTTTCCTTTTCCGATTT 19
          |||| |||| || |||||
Db      4 GGTTACCTTATCTGATTT 21


RESULT 8
5221624-13
;Patent No. 5221624
;    APPLICANT: BLAIR, LINDLEY C.;KODURI, JAR-HOW;WEICKMANN,
;JOACHIM J.
;    TITLE OF INVENTION: DNA ENCODING (LYS46, ASP97, ASP113) AND
;(LYS46, ASP113, ASP137) THAUMATIN I
;    NUMBER OF SEQUENCES: 31
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/407,416
;      FILING DATE: 14-SEP-1989
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER: 268,702
;      FILING DATE: 08-NOV-1988
;SEQ ID NO:13:
;      LENGTH: 57
5221624-13

  Query Match             66.0%;  Score 13.2;  DB 6;  Length 57;
  Best Local Similarity   83.3%;  Pred. No. 7.8e+02;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 AGGTTTCCTTTTCCGATT 18
          || ||||||||||  ||
Db     12 AGCTTTCCTTTTCCTTTT 29


RESULT 9
5221624-15
;Patent No. 5221624
;    APPLICANT: BLAIR, LINDLEY C.;KODURI, JAR-HOW;WEICKMANN,
;JOACHIM J.
;    TITLE OF INVENTION: DNA ENCODING (LYS46, ASP97, ASP113) AND
;(LYS46, ASP113, ASP137) THAUMATIN I
;    NUMBER OF SEQUENCES: 31
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/407,416
;      FILING DATE: 14-SEP-1989
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER: 268,702
;      FILING DATE: 08-NOV-1988
;SEQ ID NO:15:
;      LENGTH: 57
5221624-15

  Query Match             66.0%;  Score 13.2;  DB 6;  Length 57;
  Best Local Similarity   83.3%;  Pred. No. 7.8e+02;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 AGGTTTCCTTTTCCGATT 18
          || ||||||||||  ||
Db     12 AGCTTTCCTTTTCCTTTT 29


RESULT 10
5221624-17
;Patent No. 5221624
;    APPLICANT: BLAIR, LINDLEY C.;KODURI, JAR-HOW;WEICKMANN,
;JOACHIM J.
;    TITLE OF INVENTION: DNA ENCODING (LYS46, ASP97, ASP113) AND
;(LYS46, ASP113, ASP137) THAUMATIN I
;    NUMBER OF SEQUENCES: 31
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/407,416
;      FILING DATE: 14-SEP-1989
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER: 268,702
;      FILING DATE: 08-NOV-1988
;SEQ ID NO:17:
;      LENGTH: 57
5221624-17

  Query Match             66.0%;  Score 13.2;  DB 6;  Length 57;
  Best Local Similarity   83.3%;  Pred. No. 7.8e+02;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 AGGTTTCCTTTTCCGATT 18
          || ||||||||||  ||
Db     12 AGCTTTCCTTTTCCTTTT 29


RESULT 11
5221624-19
;Patent No. 5221624
;    APPLICANT: BLAIR, LINDLEY C.;KODURI, JAR-HOW;WEICKMANN,
;JOACHIM J.
;    TITLE OF INVENTION: DNA ENCODING (LYS46, ASP97, ASP113) AND
;(LYS46, ASP113, ASP137) THAUMATIN I
;    NUMBER OF SEQUENCES: 31
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/407,416
;      FILING DATE: 14-SEP-1989
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER: 268,702
;      FILING DATE: 08-NOV-1988
;SEQ ID NO:19:
;      LENGTH: 57
5221624-19

  Query Match             66.0%;  Score 13.2;  DB 6;  Length 57;
  Best Local Similarity   83.3%;  Pred. No. 7.8e+02;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 AGGTTTCCTTTTCCGATT 18
          || ||||||||||  ||
Db     12 AGCTTTCCTTTTCCTTTT 29


RESULT 12
US-09-603-663-74
; Sequence 74, Application US/09603663
; Patent No. 6406863
; GENERAL INFORMATION:
;  APPLICANT:  Zhu, Li
;  APPLICANT:  Hua, Shaobing
;  TITLE OF INVENTION:  HIGH THROUGHPUT GENERATION AND SCREENING OF FULLY HUMAN
;  TITLE OF INVENTION:  ANTIBODY REPERTOIRE IN YEAST
;  FILE REFERENCE:  25636-701 Seq Listing
;  CURRENT APPLICATION NUMBER: US/09/603,663
;  CURRENT FILING DATE:  2000-06-23
;  NUMBER OF SEQ ID NOS:  75
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 74
;   LENGTH: 63
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Suc 2 signal
US-09-603-663-74
```

```
  Query Match             66.0%;  Score 13.2;  DB 4;  Length 63;
  Best Local Similarity   83.3%;  Pred. No. 7.9e+02;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 AGGTTTCCTTTTCCGATT 18
          || ||||||||||  ||
Db     12 AGCTTTCCTTTTCCTTTT 29


RESULT 13
US-09-603-658-74
; Sequence 74, Application US/09603658
; Patent No. 6410246
; GENERAL INFORMATION:
;  APPLICANT:  Zhu, Li
;  APPLICANT:  Hua, Shaobing
;  TITLE OF INVENTION:  HIGHLY DIVERSE LIBRARY OF YEAST EXPRESSION VECTORS
;  FILE REFERENCE:  25636-703 Seq Listing
;  CURRENT APPLICATION NUMBER: US/09/603,658
;  CURRENT FILING DATE:  2000-06-23
;  NUMBER OF SEQ ID NOS:  75
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 74
;   LENGTH: 63
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Suc 2 signal
US-09-603-658-74

  Query Match             66.0%;  Score 13.2;  DB 4;  Length 63;
  Best Local Similarity   83.3%;  Pred. No. 7.9e+02;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 AGGTTTCCTTTTCCGATT 18
          || ||||||||||  ||
Db     12 AGCTTTCCTTTTCCTTTT 29


RESULT 14
US-09-602-373A-74
; Sequence 74, Application US/09602373A
; Patent No. 6410271
; GENERAL INFORMATION:
;  APPLICANT:  Zhu, Li
;  APPLICANT:  Hua, Shaobing B.
;  TITLE OF INVENTION:  GENERATION OF HIGHLY DIVERSE LIBRARY OF EXPRESSION
;  TITLE OF INVENTION:  VECTORS VIA HOMOLOGOUS RECOMBINATION IN YEAST
;  FILE REFERENCE:  25636-702 Seq Listing
;  CURRENT APPLICATION NUMBER: US/09/602,373A
;  CURRENT FILING DATE:  2000-06-23
;  NUMBER OF SEQ ID NOS:  76
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 74
;   LENGTH: 63
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Suc 2 signal
US-09-602-373A-74

  Query Match             66.0%;  Score 13.2;  DB 4;  Length 63;
  Best Local Similarity   83.3%;  Pred. No. 7.9e+02;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 AGGTTTCCTTTTCCGATT 18
          || ||||||||||  ||
Db     12 AGCTTTCCTTTTCCTTTT 29


RESULT 15
5221624-23
;Patent No. 5221624
;    APPLICANT: BLAIR, LINDLEY C.;KODURI, JAR-HOW;WEICKMANN,
;JOACHIM J.
;    TITLE OF INVENTION: DNA ENCODING (LYS46, ASP97, ASP113) AND
;(LYS46, ASP113, ASP137) THAUMATIN I
;    NUMBER OF SEQUENCES: 31
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/407,416
;      FILING DATE: 14-SEP-1989
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER: 268,702
;      FILING DATE: 08-NOV-1988
;SEQ ID NO:23:
;      LENGTH: 63
5221624-23

  Query Match             66.0%;  Score 13.2;  DB 6;  Length 63;
  Best Local Similarity   83.3%;  Pred. No. 7.9e+02;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 AGGTTTCCTTTTCCGATT 18
          || ||||||||||  ||
Db     18 AGCTTTCCTTTTCCTTTT 35


Search completed: November 23, 2002, 06:36:21
Job time : 22.55 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 25, 2002, 09:10:06 ; Search time 755.55 Seconds
(without alignments)
428.707 Million cell updates/sec

Title: US-09-296-264-17
Perfect score: 20
Sequence: 1 aggtttccttttccgatttc 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 357874

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	15.2	76.0	59	17	AL753463	AL753463 Arabidops
c	2	14.8	74.0	52	9	AA894524	AA894524 of90c08.s
	3	14.4	72.0	43	14	T17570	T17570 mps v30 The
c	4	14.4	72.0	55	17	AL769907	AL769907 Arabidops
c	5	14.4	72.0	83	14	BQ565295	BQ565295 gi34g10.y
	6	14	70.0	99	9	AA991706	AA991706 os92a05.s
c	7	13.8	69.0	69	17	TA203B07Q	AL476536 T. brucei
	8	13.8	69.0	80	17	BH801447	BH801447 1008116G0
c	9	13.8	69.0	96	17	BH789613	BH789613 SALK_0379
	10	13.6	68.0	30	17	AZ642644	AZ642644 1M0505F18
c	11	13.6	68.0	58	9	AI016366	AI016366 ot86d11.s
	12	13.6	68.0	63	10	AW706558	AW706558 sj59b05.y
c	13	13.6	68.0	69	17	BH790502	BH790502 SALK_0571
c	14	13.6	68.0	70	9	AI048731	AI048731 ub31e04.r
	15	13.6	68.0	78	9	AI748510	AI748510 sb54a03.y
	16	13.6	68.0	89	17	BH810855	BH810855 SALK_0512
c	17	13.6	68.0	90	14	C01576	C01576 HUMGS000858
	18	13.6	68.0	95	10	AW119558	AW119558 sd48a03.y
	19	13.6	68.0	96	10	AW119749	AW119749 sd51h08.y
	20	13.6	68.0	99	9	AU243607	AU243607 AU243607
	21	13.6	68.0	100	10	BE681958	BE681958 179807 MA
c	22	13.6	68.0	100	13	BM481766	BM481766 534252 MA
c	23	13.4	67.0	54	17	AZ920300	AZ920300 1006019C0
c	24	13.4	67.0	64	10	AW626736	AW626736 SWOvAFCAP
c	25	13.4	67.0	73	12	BG485880	BG485880 SWOv3MCAM
	26	13.4	67.0	98	17	CNS00SR2	AL088748 Arabidops
	27	13.2	66.0	46	9	AA613699	AA613699 no38f08.s
	28	13.2	66.0	52	10	AW059756	AW059756 LE3g02.yg
	29	13.2	66.0	64	9	AI130604	AI130604 SWOvL3CAN
c	30	13.2	66.0	67	17	BH847431	BH847431 SALK_0531
c	31	13.2	66.0	67	17	BH847528	BH847528 SALK_0546
	32	13.2	66.0	68	17	BH861559	BH861559 SALK_0818
c	33	13.2	66.0	70	9	AI808411	AI808411 wf94f01.x
	34	13.2	66.0	74	17	AZ818853	AZ818853 2M0089E08
c	35	13.2	66.0	76	14	T11109	T11109 hbc686 Huma
c	36	13.2	66.0	81	9	AI822095	AI822095 za75a10.x
	37	13.2	66.0	81	14	D25601	D25601 HUMGS02959
	38	13.2	66.0	91	10	BE198151	BE198151 ug75c09.y
c	39	13.2	66.0	99	13	BM121702	BM121702 L0500F06-
	40	13.2	66.0	99	17	BH893080	BH893080 3526_1_23
c	41	13	65.0	62	14	BQ567298	BQ567298 gi86d10.y
	42	13	65.0	81	17	AQ025818	AQ025818 l(2)k0712
	43	12.8	64.0	36	17	AZ371128	AZ371128 1M0122M01
	44	12.8	64.0	43	17	AZ586882	AZ586882 1M0394F11
c	45	12.8	64.0	44	17	BH814634	BH814634 SALK_0667

ALIGNMENTS

RESULT 1
AL753463
LOCUS AL753463 59 bp DNA linear GSS 17-JUN-2002
DEFINITION Arabidopsis thaliana T-DNA flanking sequence GK-050E09-012185, genomic survey sequence.
ACCESSION AL753463
VERSION AL753463.1 GI:21485961
KEYWORDS GSS.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1
AUTHORS Strizhov,N., Li,Y., Rosso,M., Viehoever,P., Dekker,K., Saedler,H. and Weisshaar,B.
TITLE A pipeline for automated high-throughput generation of FSTs (flanking sequence tags) from Arabidopsis thaliana T-DNA transformed lines
JOURNAL Unpublished
REFERENCE 2
AUTHORS Rosso,M., Strizhov,N., Li,Y., Reiss,B., Dekker,K. and Weisshaar,B.
TITLE A new Arabidopsis thaliana T-DNA mutagenised population (GABI-Kat) for flanking sequence tag based reverse genetics
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 59)
AUTHORS Li,Y., Rosso,M., Strizhov,N. and Weisshaar,B.
TITLE Direct Submission
JOURNAL Submitted (17-JUN-2002) Weisshaar B., Max-Planck-Institut fuer

```
            Zuechtungsforschung, Carl-von-Linne-Weg 10, Koeln, 50829, Germany
COMMENT     This sequence is recovered from the right border of the T-DNA. It
            indicates an insertion within the locus defined by clone T21B14.
            The sequences are generated at the MPI for Plant Breeding Research
            in the context of the GABI-Kat project. GABI-Kat is part of the
            German Plant Genomics program designated 'GABI'. Information on
            line availability can be found at:
            http://www.mpiz-koeln.mpg.de/GABI-Kat/.
FEATURES              Location/Qualifiers
     source           1. .59
                      /organism="Arabidopsis thaliana"
                      /strain="Columbia 0"
                      /db_xref="taxon:3702"
                      /clone="GK-050E09-012185"
                      /clone_lib="Arabidopsis thaliana T-DNA insertion lines"
                      /note="PCR was performed on DNA from Arabidopsis thaliana
                      plants (T1) which were transformed with the T-DNA from
                      vector pAC161. The lines contain one or more T-DNA
                      insertions. The DNA fragment(s) resulting from the PCR
                      were directly sequenced to determine the genomic sequence
                      flanking the insertion. Sequences displaying significant
                      similarity to the A. thalaina nuclear genome sequence were
                      processed for submission. T-DNA derived sequences were
                      removed"
BASE COUNT        6 a     18 c      7 g     28 t
ORIGIN

  Query Match             76.0%;  Score 15.2;  DB 17;  Length 59;
  Best Local Similarity   85.0%;  Pred. No. 1.4e+04;
  Matches   17;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 AGGTTTCCTTTTCCGATTTC 20
           | ||| ||| ||||||||||
Db       1 ATGTTCCCTATTCCGATTTC 20


RESULT 2
AA894524/c
LOCUS       AA894524                 52 bp    mRNA    linear   EST 06-APR-1998
DEFINITION  of90c08.s1 NCI_CGAP_Li5 Homo sapiens cDNA clone IMAGE:1437614 3'
            similar to TR:Q84649 Q84649 GENOME, PARTIAL SEQUENCE. ;, mRNA
            sequence.
ACCESSION   AA894524
VERSION     AA894524.1  GI:3030925
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 52)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
            R. Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html

            Trace considered overall poor quality
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 1.
FEATURES              Location/Qualifiers
     source           1. .52
                      /organism="Homo sapiens"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:1437614"
                      /clone_lib="NCI_CGAP_Li5"
                      /tissue_type="hepatic adenoma"
                      /lab_host="DH10B"
                      /note="Organ: liver; Vector: pCMV-SPORT4; Site_1: SalI;
                      Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                      Average insert size 0.8 kb."
BASE COUNT       26 a      4 c     22 g      0 t
ORIGIN

  Query Match             74.0%;  Score 14.8;  DB 9;  Length 52;
  Best Local Similarity   88.9%;  Pred. No. 2.1e+04;
  Matches   16;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       3 GTTTCCTTTTCCGATTTC 20
           | ||||||||||| ||||
Db      32 GCTTCCTTTTCCGCTTTC 15


RESULT 3
T17570
LOCUS       T17570                   43 bp    mRNA    linear   EST 06-JUN-1994
DEFINITION  mps v30 The blue guys library Saccharomyces cerevisiae cDNA
            sequence upstream of LacZ fusion similar to RNA12, S92205, mRNA
            sequence.
ACCESSION   T17570
VERSION     T17570.1  GI:458592
KEYWORDS    EST.
SOURCE      baker's yeast.
  ORGANISM  Saccharomyces cerevisiae
            Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
            Saccharomycetales; Saccharomycetaceae; Saccharomyces.
REFERENCE   1  (bases 1 to 43)
  AUTHORS   Burns,N., Grimwade,B., Ross-Macdonald,P.B., Choi,E.-Y., Finberg,K.,
            Roeder,G.S. and Snyder,M.
  TITLE     Large-scale analysis of gene expression, protein localization and
            gene disruption in Saccharomyces cerevisiae
  JOURNAL   Genes Dev. 8, 1087-1105 (1994)
  MEDLINE   95011603
COMMENT     Contact: Snyder M
            Department of Biology
            Yale University
            New Haven CT 06520-8103
            Tel: 2034326139
            Fax: 2034326161
            Email: snymicp@yalevm.ycc.yale.edu
            LacZ fusion; Vegetative expression; Beta-gal fusion localization
            pattern: mitochondrial; Disruption phenotype: none detected;
            Fusion: codon 626 of RNA12 gene.  Sequence below near or adjacent
            to lacZ.
            Seq primer: LacZ sequences in transposon.
FEATURES              Location/Qualifiers
     source           1. .43
                      /organism="Saccharomyces cerevisiae"
                      /db_xref="taxon:4932"
                      /clone_lib="The blue guys library"
                      /lab_host="E.coli"
                      /note="Vector: pRECmTn; A yeast genomic DNA library was
                      prepared in the vector pHSS6, and subjected to transposon
                      mutagenesis with mTn3. This mini-transposon carries lacZ
                      sequences that lack an initiation codon; expression of
                      lacZ is only provided by in frame fusion to yeast coding
                      sequence. The yeast genomic DNA carrying the transposon
                      was excised from pHSS6 and transplaced back onto the yeast
                      chromosome. Yeast colonies expressing lacZ were screened
                      for in a color assay. A plasmid containing the genomic
                      DNA/lacZ fusion junction was recovered from each
                      individual yeast colony that expressed lacZ activity.
                      These recovered plasmids comprise 'The blue guys library'.
                      The fusion junction was then sequenced to identify the
                      expressed ORF upstream of the fusion. "
BASE COUNT        8 a      8 c      5 g     22 t
ORIGIN
```

```
  Query Match           72.0%;  Score 14.4;  DB 14;  Length 43;
  Best Local Similarity  93.8%;  Pred. No. 3.2e+04;
  Matches   15;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      3 GTTTCCTTTTCCGATT 18
          ||||||||||| ||||
Db     18 GTTTCCTTTTCTGATT 33


RESULT 4
AL769907/c
LOCUS       AL769907                 55 bp    DNA     linear   GSS 19-JUN-2002
DEFINITION  Arabidopsis thaliana T-DNA flanking sequence GK-094D08-012026,
            genomic survey sequence.
ACCESSION   AL769907
VERSION     AL769907.1  GI:21532109
KEYWORDS    GSS.
SOURCE      thale cress.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1
  AUTHORS   Strizhov,N., Li,Y., Rosso,M., Viehoever,P., Dekker,K., Saedler,H.
            and Weisshaar,B.
  TITLE     A pipeline for automated high-throughput generation of FSTs
            (flanking sequence tags) from Arabidopsis thaliana T-DNA
            transformed lines
  JOURNAL   Unpublished
REFERENCE   2
  AUTHORS   Rosso,M., Strizhov,N., Li,Y., Reiss,B., Dekker,K. and Weisshaar,B.
  TITLE     A new Arabidopsis thaliana T-DNA mutagenised population (GABI-Kat)
            for flanking sequence tag based reverse genetics
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 55)
  AUTHORS   Li,Y., Strizhov,N., Rosso,M. and Weisshaar,B.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JUN-2002) Weisshaar B., Max-Planck-Institut fuer
            Zuechtungsforschung, Carl-von-Linne-Weg 10, Koeln, 50829, Germany
COMMENT     This sequence is recovered from the left border of the T-DNA. It
            indicates an insertion close to or within gene At3g20430. The
            sequences are generated at the MPI for Plant Breeding Research in
            the context of the GABI-Kat project. GABI-Kat is part of the German
            Plant Genomics program designated 'GABI'. Information on line
            availability can be found at:
            http://www.mpiz-koeln.mpg.de/GABI-Kat/.
FEATURES             Location/Qualifiers
     source          1. .55
                     /organism="Arabidopsis thaliana"
                     /strain="Columbia 0"
                     /db_xref="taxon:3702"
                     /clone="GK-094D08-012026"
                     /clone_lib="Arabidopsis thaliana T-DNA insertion lines"
                     /note="PCR was performed on DNA from Arabidopsis thaliana
                     plants (T1) which were transformed with the T-DNA from
                     vector pAC161. The lines contain one or more T-DNA
                     insertions. The DNA fragment(s) resulting from the PCR
                     were directly sequenced to determine the genomic sequence
                     flanking the insertion. Sequences displaying significant
                     similarity to the A. thalaina nuclear genome sequence were
                     processed for submission. T-DNA derived sequences were
                     removed"
BASE COUNT       21 a     10 c     15 g      9 t
ORIGIN

  Query Match           72.0%;  Score 14.4;  DB 17;  Length 55;
  Best Local Similarity  93.8%;  Pred. No. 3.2e+04;
  Matches   15;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      4 TTTCCTTTTCCGATTT 19
          |||||||||||| |||
Db     52 TTTCCTTTTCCGGTTT 37


RESULT 5
BQ565295/c
LOCUS       BQ565295                 83 bp    mRNA    linear   EST 19-JUN-2002
DEFINITION  gi34g10.y1 Mouse Organ of Corti cDNA pBluescript Mus musculus cDNA
            clone gi34g10 5', mRNA sequence.
ACCESSION   BQ565295
VERSION     BQ565295.1  GI:21468612
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 83)
  AUTHORS   Kachar,B.
  TITLE     EST analysis of gene expression in the mouse Organ of Corti at the
            onset of hearing
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Kachar,B.
            Structural Cell Biology
            National Institute of Deafness and other Communication Disorders
            50/4249 South Drive, NIH, Bethesda, MD 20892-8027, USA
            Tel: 301-402-1599
            Fax: 301-402-1765
            Email: kacharb@nidcd.nih.gov
            Plate: 34  row: g  column: 10
            Seq primer: M13RP1 reverse primer (ABI).
FEATURES             Location/Qualifiers
     source          1. .83
                     /organism="Mus musculus"
                     /strain="BALB/c"
                     /db_xref="taxon:10090"
                     /clone="gi34g10"
                     /clone_lib="Mouse Organ of Corti cDNA pBluescript"
                     /sex="male and female"
                     /dev_stage="Post natal day 5 to 13"
                     /note="Organ: Organ of Corti; Vector: pBluescript; The
                     organ of Corti (OC) was fine dissected from a total of 386
                     OC as follows: 102 samples from post-natal (P) day 5; 72
                     from P6; 60 from P7; 46 from P8; 18 from P9; 20 from P10;
                     14 from P12 and 24 from P13. After killing animals by
                     cervical dislocation followed by decapitation, the bulla
                     was removed and opened in Leibowitz medium. The bony
                     capsule of the cochlea was chipped away, stria vascularis
                     and spiral ligament were removed and the sensory
                     epithelium was carefully dissected out of the modiolus.
                     Total RNA was extracted using the micro Fasttrack kit
                     (catalog # K1593-02; Invitrogen, Carlsbad, CA), according
                     to manufacturer's instructions. Reverse transcription and
                     library construction were carried out with the Uni-Zap XR
                     vector kit (catalog # 237211, Stratagene) and Uni-Zap XR
                     Gigapack III Gold Cloning kit (catalog # 237612), both
                     from Stratagene (La Jolla, CA, USA), according to
                     manufacturer's instructions. Briefly: 1.5 ug mRNA was
                     reverse transcribed using a hybrid oligo(dT) linker-primer
                     that contains an Xho I site. First strand synthesis was
                     primed with the linker- primer and transcribed using
                     Moloney murine leukemia virus reverse transcriptase
                     (MMLV-RT) and 5-methyl dCTP. The second strand was
                     synthesized with DNA polymerase and RNase H. Complementary
                     DNA was blunt ended with Pfu DNA polymerase, ligated with
                     EcoR I adapters in the presence of ligase and digested
                     with Xho I. The cDNA was sequentially size fractionated
                     over Pharmacia Size Sep400 (Pharmacia, Uppsala, Sweden)
                     and Clontech Chroma Spin-1000 (Clontech, Palo Alto, CA)
                     columns to enrich for cDNAs greater than 400bp and 1000 bp
                     , respectively. The cDNA was then directionally ligated to
                     the Uni-ZAP XR vector, which had been predigested with
                     EcoR I and Xho I. The phagemid was packaged with Gigapak
                     III Gold and, upon titration on XL1 Blue MRF' cells, the
                     yield of the phage library was estimated to be 11,100,000
                     recombinants. Stratagene's ExAssist Interference
                     resistance helper phage (catalogue # 211203) was adopted
```

to rescue plasmid DNA from the phages. Upon plating of the rescued library, individual cDNA clones were selected and grown in 96-well, 2 ml growth plate. Plasmid DNA was purified from 200 ul of saturated culture with the Concert96(TM) plasmid purification kit (Invitrogen, Carlsbad, CA) as instructed by the manufacturer. ESTs from the 5' end of the cDNA clones were generated with the universal M13 reverse primer (CAGGAAACAGCTATGACC) and 25% strength BigDye terminator sequencing chemistry (Applied Biosystems, Foster City, CA). Sequencing reactions were performed on MJ Tetrad thermal cyclers (MJ Research, Waltham, MA), and analyzed on 3700 automated capillary sequencers using POP5 polymer (Applied Biosystems, Foster City, CA). The frequency distribution of the library is as follows: 72% of genes have 1 copy; 14.3% 2; 12% 3-10; 1.4% 11-50 and 0.1% 51-150. As to gene function, 45% of genes are present in GenBank and have know function; 23% have hits in GenBank, but do not have assigned function; 12% are uncharacterized ESTs and 20% are unidentified."

```
BASE COUNT       29 a     12 c     23 g     19 t
ORIGIN

  Query Match             72.0%;  Score 14.4;  DB 14;  Length 83;
  Best Local Similarity   93.8%;  Pred. No. 3.2e+04;
  Matches   15;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       4 TTTCCTTTTCCGATTT 19
           ||| ||||||||||||
Db      56 TTTGCTTTTCCGATTT 41


RESULT 6
AA991706
LOCUS       AA991706                 99 bp    mRNA    linear   EST 03-JUN-1998
DEFINITION  os92a05.s1 NCI_CGAP_GC3 Homo sapiens cDNA clone IMAGE:1612784 3'
            similar to gb:L05095 60S RIBOSOMAL PROTEIN L30 (HUMAN);contains
            PTR5.t3 TAR1 repetitive element ;, mRNA sequence.
ACCESSION   AA991706
VERSION     AA991706.1  GI:3178195
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 99)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html

            Trace considered overall poor quality
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .99
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1612784"
                     /clone_lib="NCI_CGAP_GC3"
                     /tissue_type="pooled germ cell tumors"
                     /lab_host="DH10B"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; 1st strand cDNA was prepared from 3 pooled
                     germ cell tumors, and was then primed with a Not I -
                     oligo(dT) primer. Double-stranded cDNA was ligated to Eco
                     RI adaptors (Pharmacia), digested with Not I and cloned
                     into the Not I and Eco RI sites of the modified pT7T3
                     vector.  Library is not normalized. Library was
                     constructed by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT       22 a     15 c     25 g     37 t
ORIGIN

  Query Match             70.0%;  Score 14;  DB 9;  Length 99;
  Best Local Similarity   100.0%;  Pred. No. 4.8e+04;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 AGGTTTCCTTTTCC 14
           ||||||||||||||
Db      70 AGGTTTCCTTTTCC 83


RESULT 7
TA203B07Q/c
LOCUS       TA203B07Q                69 bp    DNA     linear   GSS 13-DEC-2000
DEFINITION  T. brucei sheared genomic DNA clone 203b07, reverse sequence,
            genomic survey sequence.
ACCESSION   AL476536
VERSION     AL476536.1  GI:11843203
KEYWORDS    GSS.
SOURCE      Trypanosoma brucei.
  ORGANISM  Trypanosoma brucei
            Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
            Trypanosoma.
REFERENCE   1  (bases 1 to 69)
  AUTHORS   Hall,N., Bowman,S., Lennard,N.J., Doggett,J., Atkin,R.,
            Chillingworth,C., Ormond,D., Harris,B., El-Sayed,N., Hou,L.,
            Melville,S.E., Rajandream,M.A. and Barrell,B.G.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-DEC-2000) Trypanosoma brucei genome sequencing
            project, Sanger Centre, The Wellcome Trust Genome Campus, Hinxton,
            Cambridge CB10 1SA, E-mail: barrell@sanger.ac.uk and
            nh1@sanger.ac.uk
COMMENT     Constructed at the Institute for Genomic Research (TIGR),
            Rockville, MD. Genomic DNA isolated from a cloned population of
            Trypanosoma brucei (TREU927/4 GUTat 10.1) was mechanically sheared
            to give a tight size distribution (
            4 kb).  The v + i method used for the library construction is
            described in detail in Smith, H. and Venter, J.C. (Making small
            insert libraries for whole genome shotgun sequencing projects.  In
            Genome Sequencing: A Practical Approach, eds. M. Vaudin and B.
            Barrell, Oxford University Press, 1999).
            Email: nelsayed@tigr.org
            Details of T. brucei sequencing at the Sanger Centre are available
            at http://www.sanger.ac.uk/Projects/T_brucei/.
FEATURES             Location/Qualifiers
     source          1. .69
                     /organism="Trypanosoma brucei"
                     /strain="TREU927"
                     /db_xref="taxon:5691"
                     /clone="203b07"
BASE COUNT       36 a      5 c     20 g      8 t
ORIGIN

  Query Match             69.0%;  Score 13.8;  DB 17;  Length 69;
  Best Local Similarity   88.2%;  Pred. No. 5.7e+04;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       4 TTTCCTTTTCCGATTTC 20
           ||| ||||||| |||||
Db      25 TTTTCTTTTCCCATTTC 9


RESULT 8
BH801447
LOCUS       BH801447                 80 bp    DNA     linear   GSS 25-APR-2002
DEFINITION  1008116G02.1EL_x1 1008 - RescueMu Grid I Zea mays genomic, DNA
```

```
            sequence.
ACCESSION   BH801447
VERSION     BH801447.1  GI:20314525
KEYWORDS    GSS.
SOURCE      Zea mays.
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 80)
  AUTHORS   Walbot,V.
  TITLE     Maize genomic sequences found using engineered RescueMu transposon
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Walbot V
            Department of Biological Sciences
            Stanford University
            855 California Ave, Palo Alto, CA 94304, USA
            Tel: 650 723 2227
            Fax: 650 725 8221
            Email: walbot@stanford.edu
            Very probable ligation site of ends cut by single endonuclease.
            Reverse complemented post-ligation sequence from source sequence.
            Plate: 1008116  row: 16
            Class: transposon-tagged.
FEATURES             Location/Qualifiers
     source          1. .80
                     /organism="Zea mays"
                     /cultivar="mixed background W23/A188/B73"
                     /db_xref="taxon:4577"
                     /clone_lib="1008 - RescueMu Grid I"
                     /tissue_type="leaf"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /note="Organ: leaf; Vector: RescueMu (engineered from
                     pBlueScript backbone); Site_1: BamHI; Site_2: BglII;
                     RescueMu is a 4.9 kb, modified maize Mu transposon
                     designed to allow plasmid rescue from total genomic DNA.
                     Mu elements insert preferentially into transcription
                     units. For more information on RescueMu, go to the  web
                     site www.zmdb.iastate.edu and follow the links for
                     'RescueMu.' Grid I was grown at Berkeley in 2001. DNA was
                     extracted from leaf punches,  double digested using BamHI
                     and BglII, and ligated to form circular plasmids.  DH10B
                     cells were transformed and then screened on LB plates with
                     ampicillin."
BASE COUNT       17 a      22 c      16 g      24 t       1 others
ORIGIN

  Query Match             69.0%;  Score 13.8;  DB 17;  Length 80;
  Best Local Similarity   88.2%;  Pred. No. 5.8e+04;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       4 TTTCCTTTTCCGATTTC 20
           ||||||||||  ||||
Db      16 TTTCCTTTTCCTCTTTC 32


RESULT 9
BH789613/c
LOCUS       BH789613                  96 bp    DNA     linear   GSS 02-APR-2002
DEFINITION  SALK_037911.47.65.x Arabidopsis thaliana TDNA insertion lines
            Arabidopsis thaliana genomic clone SALK_037911.47.65.x, DNA
            sequence.
ACCESSION   BH789613
VERSION     BH789613.1  GI:19882711
KEYWORDS    GSS.
SOURCE      thale cress.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 96)
  AUTHORS   Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R., Gadrinab
            ,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L., Shinn,P.
            , Zimmerman,J. and Ecker,J.R.
  TITLE     A Sequence-Indexed Library of Insertion Mutations in the
            Arabidopsis Genome
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Joseph R. Ecker
            Salk Institute Genomic Analysis Laboratory (SIGnAL)
            The Salk Institute for Biological Studies
            10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
            Tel: 858 453 4100 x1752
            Fax: 858 558 6379
            Email: ecker@salk.edu
            This is single pass sequence recovered from the left border of
            TDNA.
            Class: TDNA tagged.
FEATURES             Location/Qualifiers
     source          1. .96
                     /organism="Arabidopsis thaliana"
                     /strain="Columbia 0"
                     /db_xref="taxon:3702"
                     /clone="SALK_037911.47.65.x"
                     /clone_lib="Arabidopsis thaliana TDNA insertion lines"
                     /note="PCR was performed on Arabidopsis thaliana lines
                     each of which contains  one or more TDNA insertion
                     elements. The resultant fragment for each  line was
                     directly sequenced to determine the genomic sequence at
                     the  site of insertion. Details of the protocols used can
                     be found at  http://signal.salk.edu/tdna_protocols.html"
BASE COUNT       48 a      8 c      18 g      22 t
ORIGIN

  Query Match             69.0%;  Score 13.8;  DB 17;  Length 96;
  Best Local Similarity   88.2%;  Pred. No. 5.8e+04;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       4 TTTCCTTTTCCGATTTC 20
           ||||||  ||||  |||||
Db      78 TTTCCTCTTCCAATTTC 62


RESULT 10
AZ642644
LOCUS       AZ642644                  30 bp    DNA     linear   GSS 14-DEC-2000
DEFINITION  1M0505F18R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0505F18 R, DNA sequence.
ACCESSION   AZ642644
VERSION     AZ642644.1  GI:11769456
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 30)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0505  row: F  column: 18
            Seq primer: CACACAGGAAACAGCTATGACC
            Class: plasmid ends
            High quality sequence stop: 30.
```

```
FEATURES             Location/Qualifiers
     source          1. .30
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC1M0505F18"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
BASE COUNT        5 a      5 c      5 g     15 t
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 17;  Length 30;
  Best Local Similarity   80.0%;  Pred. No. 6.9e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 AGGTTTCCTTTTCCGATTTC 20
           |||| | |||||| |||| |
Db       5 AGGTGTTCTTTTCTGATTAC 24


RESULT 11
AI016366/c
LOCUS       AI016366                 58 bp    mRNA    linear   EST 27-AUG-1998
DEFINITION  ot86d11.s1 Soares_total_fetus_Nb2HF8_9w Homo sapiens cDNA clone
            IMAGE:1623669 3' similar to TR:O00410 O00410 KARYOPHERIN BETA 3. ;,
            mRNA sequence.
ACCESSION   AI016366
VERSION     AI016366.1  GI:3230702
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 58)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Trace considered overall poor quality
            Insert Length: 724   Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .58
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1623669"
                     /clone_lib="Soares_total_fetus_Nb2HF8_9w"
                     /dev_stage="8-9 weeks"
                     /lab_host="DH10B"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                     was prepared from mRNA obtained from pooled 8-9 week
                     (total) fetus material with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCTTAATTTTTTTTTTTTTTTT 3']. 
                     Double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. Library
                     went through one round of normalization, and was
                     constructed  by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT       22 a     15 c     13 g      8 t
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 9;  Length 58;
  Best Local Similarity   80.0%;  Pred. No. 7e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 AGGTTTCCTTTTCCGATTTC 20
           |||||||||| | ||  ||||
Db      23 AGGTTTCCTCTGCCTGTTTC 4


RESULT 12
AW706558
LOCUS       AW706558                 63 bp    mRNA    linear   EST 03-DEC-2001
DEFINITION  sj59b05.y1 Gm-c1033 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
            Gm-c1033-1546 5', mRNA sequence.
ACCESSION   AW706558
VERSION     AW706558.1  GI:7590814
KEYWORDS    EST.
SOURCE      soybean.
  ORGANISM  Glycine max
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
            Glycine.
REFERENCE   1  (bases 1 to 63)
  AUTHORS   Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Coryell,V., Khanna
            ,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
            Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers
            ,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk
            ,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann
            ,R., Waterston,R. and Wilson,R.
  TITLE     Public Soybean EST Project
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Shoemaker R/Public Soybean EST Project
            Public Soybean EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available through: ResGen, Invitrogen Corp. 2130
            South Memorial Parkway Huntsville, AL 35801 For further information
            call: (800)-533-4363 or contact via email: ccu@resgen.com
            High quality sequence stop: 58.
FEATURES             Location/Qualifiers
     source          1. .63
                     /organism="Glycine max"
                     /db_xref="taxon:3847"
                     /clone="GENOME SYSTEMS CLONE ID: Gm-c1033-1546"
                     /clone_lib="Gm-c1033"
                     /tissue_type="'Desloy 5710' seedling roots"
                     /lab_host="DH10B"
                     /note="Vector: pBluescript II XR; Site_1: EcoRI; Site_2:
                     XhoI; This cDNA library was constructed from mRNA isolated
                     from 'Desloy 5710' seedling roots. Tissue was taken from
                     7-day-old seedlings that had been propagated on paper
                     towels with distilled water. Tissue was taken from the tip
                     to the first lateral root, usually about 3cm from the tip,
                     and flash-frozen in liquid nitrogen. Stratagene's cDNA
                     Synthesis Kit (catalog number 200401) was used to
```

synthesize the cDNA. First-strand synthesis was performed with 5-methyl dCTP, hence the ligated cDNA was hemimethylated. A modification of Stratagene's first-strand synthesis primer was used. An 'anchor' nucleotide (V=A, C, or G) was added to the 3' end of the primer [GAGAGAGAGAGAGAGAGAACTAGTCTCGAG(T)18V] to anchor the primer at the 5' end of the poly(A) tract. After second strand synthesis, the cDNA ends were filled in with cloned Pfu DNA, ligated to EcoRI adapters and subsequently phosphorylated. The cDNA was then precipitated and redissolved in sterile, RNase-, DNase-free water. The XhoI site within the first-strand synthesis primer was then restricted by digestion with XhoI from Promega (40U/ul); all XhoI sites in the cDNA would be protected by their hemimethylated status. The cDNA constructs were size-fractionated with a 500 bp cutoff, using GibcoBRL Life Technologies' cDNA Size Fractionation column. The column eluent was then precipitated, redissolved, and ligated into Stratagene's pBluescript II XR Predigested vector (pBluescript II SK(+) vector that has been digested with EcoRI and XhoI, and phosphorylated by Stratagene). 100% of the white and blue colonies appear to contain recombinant plasmids with cDNA inserts, based on size (n=24 and 6, respectively). This library was constructed by Dr. Paul Keim and Dr. Virginia Coryell."

BASE COUNT 23 a 14 c 12 g 14 t
ORIGIN

Query Match 68.0%; Score 13.6; DB 10; Length 63;
Best Local Similarity 80.0%; Pred. No. 7e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
Qy       1 AGGTTTCCTTTTCCGATTTC 20
           |||||||| || || | |||
Db      35 AGGTTTCCATTGCCAAGTTC 54
```

RESULT 13
BH790502/c
LOCUS BH790502 69 bp DNA linear GSS 02-APR-2002
DEFINITION SALK_057190.20.70.x Arabidopsis thaliana TDNA insertion lines Arabidopsis thaliana genomic clone SALK_057190.20.70.x, DNA sequence.
ACCESSION BH790502
VERSION BH790502.1 GI:19883600
KEYWORDS GSS.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 69)
AUTHORS Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R., Gadrinab ,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L., Shinn,P. , Zimmerman,J. and Ecker,J.R.
TITLE A Sequence-Indexed Library of Insertion Mutations in the Arabidopsis Genome
JOURNAL Unpublished (2001)
COMMENT Contact: Joseph R. Ecker
Salk Institute Genomic Analysis Laboratory (SIGnAL)
The Salk Institute for Biological Studies
10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
Tel: 858 453 4100 x1752
Fax: 858 558 6379
Email: ecker@salk.edu
This is single pass sequence recovered from the left border of TDNA. This sequence lies within an annotated exon of At1g68640.
Class: TDNA tagged.
FEATURES Location/Qualifiers
source 1. .69
/organism="Arabidopsis thaliana"
/strain="Columbia 0"
/db_xref="taxon:3702"
/clone="SALK_057190.20.70.x"
/clone_lib="Arabidopsis thaliana TDNA insertion lines"
/note="PCR was performed on Arabidopsis thaliana lines each of which contains one or more TDNA insertion elements. The resultant fragment for each line was directly sequenced to determine the genomic sequence at the site of insertion. Details of the protocols used can be found at http://signal.salk.edu/tdna_protocols.html"
BASE COUNT 24 a 13 c 16 g 16 t
ORIGIN

Query Match 68.0%; Score 13.6; DB 17; Length 69;
Best Local Similarity 80.0%; Pred. No. 7e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
Qy       1 AGGTTTCCTTTTCCGATTTC 20
           ||||| | | ||| ||||||
Db      45 AGGTTCCATGTTCTGATTTC 26
```

RESULT 14
AI048731/c
LOCUS AI048731 70 bp mRNA linear EST 08-JUL-1998
DEFINITION ub31e04.r1 Soares_thymus_2NbMT Mus musculus cDNA clone IMAGE:1379358 5' similar to TR:O35309 O35309 NMI. ;, mRNA sequence.
ACCESSION AI048731
VERSION AI048731.1 GI:3297018
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 70)
AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.
TITLE The WashU-HHMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Marra M/Mouse EST Project
WashU-HHMI Mouse EST Project
Washington University School of MedicineP
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:901826
Trace considered overall poor quality
Possible reversed clone: similarity on wrong strand
Seq primer: -28m13 rev2 ET from Amersham
High quality sequence stop: 1.
FEATURES Location/Qualifiers
source 1. .70
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:1379358"
/clone_lib="Soares_thymus_2NbMT"
/sex="male"
/tissue_type="Thymus"
/dev_stage="4 weeks"
/lab_host="DH10B"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGGAGCGGCCGCGTTTTTTTTTTTTTTTTTTTTTTTT 3']; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. RNA

provided by Dr. Bertrand Jordan. Library went through two rounds of normalization, and was constructed by Bento Soares and M.Fatima Bonaldo."
BASE COUNT 25 a 13 c 14 g 18 t
ORIGIN

Query Match 68.0%; Score 13.6; DB 9; Length 70;
Best Local Similarity 80.0%; Pred. No. 7e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
Qy       1 AGGTTTCCTTTTCCGATTTC 20
           |||||| ||||| | | |||
Db      32 AGGTTTTCTTTTTCAAATTC 13
```

RESULT 15
AI748510
LOCUS AI748510 78 bp mRNA linear EST 30-NOV-2001
DEFINITION sb54a03.y1 Gm-c1016 Glycine max cDNA clone GENOME SYSTEMS CLONE ID: Gm-c1016-293 5', mRNA sequence.
ACCESSION AI748510
VERSION AI748510.1 GI:5126774
KEYWORDS EST.
SOURCE soybean.
ORGANISM Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
REFERENCE 1 (bases 1 to 78)
AUTHORS Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Coryell,V., Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.
TITLE Public Soybean EST Project
JOURNAL Unpublished (1999)
COMMENT Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: ResGen, Invitrogen Corp. 2130 South Memorial Parkway Huntsville, AL 35801 For further information call: (800)-533-4363 or contact via email: ccu@resgen.com.
FEATURES Location/Qualifiers
source 1. .78
/organism="Glycine max"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-c1016-293"
/clone_lib="Gm-c1016"
/tissue_type="immature flowers of field grown plants"
/lab_host="XL10-Gold"
/note="Vector: pBluescript II XR; Site_1: EcoRI; Site_2: XhoI; This cDNA library was constructed from mRNA isolated from immature flowers of field grown plants. The cDNA library was prepared using the Stratagene pBluescript II XR library construction kit. Complementary DNA was synthesized from mRNA using a primer consisting of a poly (dT) sequence with a XhoI restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into XL10-Gold host cells. This library was constructed by Dr. Randy Shoemaker and Dr. John Erpelding."
BASE COUNT 27 a 15 c 9 g 27 t
ORIGIN

Query Match 68.0%; Score 13.6; DB 9; Length 78;
Best Local Similarity 80.0%; Pred. No. 7e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
Qy       1 AGGTTTCCTTTTCCGATTTC 20
           ||  |||||||||  ||||
Db      13 AGAATTCCTTTTCCTTTTC 32
```

Search completed: November 26, 2002, 04:09:08
Job time : 765.8 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 3, 2002, 14:46:40 ; Search time 302.2 Seconds
(without alignments)
1926.063 Million cell updates/sec

Title: US-09-296-264-18
Perfect score: 20
Sequence: 1 gtgctccctgtttcatcaat 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 995600

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
c 1	14.2	71.0	90	8	AF064262	AF064262 Saccharom
2	13.4	67.0	42	9	S63149	S63149 hprt=hypoxa
3	13.4	67.0	93	6	AR111310	AR111310 Sequence
c 4	13.2	66.0	49	6	AR123848	AR123848 Sequence
c 5	13.2	66.0	57	6	AX103651	AX103651 Sequence
c 6	13.2	66.0	61	1	ECO5ERNAC	M16642 E.coli 5'-t
7	13	65.0	21	6	AX402155	AX402155 Sequence
c 8	13	65.0	27	6	AR039988	AR039988 Sequence
c 9	13	65.0	27	6	AR189295	AR189295 Sequence
c 10	12.8	64.0	30	6	AR070500	AR070500 Sequence
c 11	12.8	64.0	30	6	AX128522	AX128522 Sequence
c 12	12.8	64.0	30	6	E26044	E26044 Peptide fra
c 13	12.8	64.0	31	6	AR195964	AR195964 Sequence
c 14	12.8	64.0	43	6	AX108135	AX108135 Sequence
c 15	12.8	64.0	84	9	AF274789	AF274789 Homo sapi
c 16	12.6	63.0	20	6	AX378657	AX378657 Sequence
c 17	12.6	63.0	20	6	AX462655	AX462655 Sequence
c 18	12.6	63.0	24	6	AX393018	AX393018 Sequence
c 19	12.6	63.0	30	6	I34500	I34500 Sequence 9
c 20	12.6	63.0	30	6	I57335	I57335 Sequence 9
c 21	12.6	63.0	30	6	I73213	I73213 Sequence 9
22	12.6	63.0	34	6	AX474364	AX474364 Sequence
23	12.6	63.0	36	6	AX474362	AX474362 Sequence
24	12.6	63.0	37	6	AR073767	AR073767 Sequence
25	12.6	63.0	37	6	AR208263	AR208263 Sequence
26	12.6	63.0	38	6	AR096264	AR096264 Sequence
27	12.6	63.0	41	6	A46386	A46386 Sequence 8
28	12.6	63.0	41	6	AR078962	AR078962 Sequence
c 29	12.6	63.0	45	6	AX175593	AX175593 Sequence
30	12.6	63.0	63	6	AR073782	AR073782 Sequence
31	12.6	63.0	63	6	AR208278	AR208278 Sequence
32	12.6	63.0	65	6	AX483988	AX483988 Sequence
c 33	12.6	63.0	90	6	E05553	E05553 cDNA encodi
c 34	12.6	63.0	90	6	E05554	E05554 cDNA encodi
c 35	12.6	63.0	100	11	AB059124	AB059124 Sus scrof
36	12.4	62.0	16	6	A14412	A14412 oligonucleo
37	12.4	62.0	25	6	AX118564	AX118564 Sequence
c 38	12.4	62.0	25	6	AX196730	AX196730 Sequence
c 39	12.4	62.0	27	6	AR039608	AR039608 Sequence
c 40	12.4	62.0	27	6	AR184939	AR184939 Sequence
c 41	12.4	62.0	27	6	AR185181	AR185181 Sequence
c 42	12.4	62.0	27	6	AR185576	AR185576 Sequence
c 43	12.4	62.0	27	6	AR191141	AR191141 Sequence
c 44	12.4	62.0	27	6	AR191320	AR191320 Sequence
c 45	12.4	62.0	27	6	AR191591	AR191591 Sequence

ALIGNMENTS

RESULT 1
AF064262/c
LOCUS AF064262 90 bp snoRNA linear PLN 16-OCT-2001
DEFINITION Saccharomyces cerevisiae snR50 small nucleolar RNA, complete sequence.
ACCESSION AF064262
VERSION AF064262.1 GI:3135699
KEYWORDS .
SOURCE Saccharomyces cerevisiae.
ORGANISM Saccharomyces cerevisiae
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; Saccharomycetales; Saccharomycetaceae; Saccharomyces.
REFERENCE 1 (bases 1 to 90)
AUTHORS Lowe,T.M. and Eddy,S.R.
TITLE Identification and verification of 22 new 2'-O-methyl guide snoRNAs

```
            in yeast
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 90)
  AUTHORS   Ni,J. and Fournier,M.J.
  JOURNAL   Unpublished
  REMARK    Independent identification as C/D box snoRNA
REFERENCE   3  (bases 1 to 90)
  AUTHORS   Lowe,T.M. and Eddy,S.R.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-MAY-1998) Dept. Genetics, Washington University
            School of Medicine, 660 S. Euclid, St. Louis, MO 63110, USA
COMMENT     Gene function confirmed by homologous gene disruption.
            Mature snoRNA 5' end determined experimentally (+/-) 1 nt; mature
            snoRNA 3' end computationally predicted.
            Experimental evidence of methyl guide function available at
            'http://rna.wustl.edu/snoRNAdb/'.
FEATURES             Location/Qualifiers
     source          1. .90
                     /organism="Saccharomyces cerevisiae"
                     /db_xref="taxon:4932"
                     /chromosome="XV"
                     /map="between YOL034W and YOL035C"
     gene            1. .90
                     /gene="SNR50"
                     /note="Z14 snoRNA"
     snoRNA          1. .90
                     /gene="SNR50"
                     /product="snR50 small nucleolar RNA"
                     /note="guides 2'-O-methylation on large subunit rRNA at
                     G865"
                     /evidence=experimental
     misc_feature    6. .12
                     /gene="SNR50"
                     /note="box C"
     misc_feature    72. .83
                     /gene="SNR50"
                     /note="large subunit ribosomal RNA complementarity"
     misc_feature    85. .88
                     /gene="SNR50"
                     /note="box D"
BASE COUNT       34 a     14 c     17 g     25 t
ORIGIN

  Query Match             71.0%;  Score 14.2;  DB 8;  Length 90;
  Best Local Similarity   84.2%;  Pred. No. 1.3e+04;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 TGCTCCCTGTTTCATCAAT 20
          |||| ||| |||||||| |
Db     23 TGCTACCTCTTTCATCATT 5


RESULT 2
S63149
LOCUS       S63149                    42 bp    DNA     linear   PRI 07-MAY-1993
DEFINITION  hprt=hypoxanthine-guanine phosphoribosyltransferase [human, fetal
            T-lymphocytes, Genomic Mutant, 42 nt].
ACCESSION   S63149
VERSION     S63149.1  GI:237842
KEYWORDS    .
SOURCE      Homo sapiens fetal T-lymphocytes.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 42)
  AUTHORS   Fuscoe,J.C., Zimmerman,L.J., Lippert,M.J., Nicklas,J.A.,
            O'Neill,J.P. and Albertini,R.J.
  TITLE     V(D)J recombinase-like activity mediates hprt gene deletion in
            human fetal T-lymphocytes
  JOURNAL   Cancer Res. 51 (21), 6001-6005 (1991)
  MEDLINE   92034710
   PUBMED   1933863
  REMARK    GenBank staff at the National Library of Medicine created this
            entry [NCBI gibbsq 63149] from the original journal article.
            This sequence comes from Fig.2A.
COMMENT     breakpoint junctions of deletion mutations.
FEATURES             Location/Qualifiers
     source          1. .42
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     gene            1. .42
                     /partial
                     /gene="hprt"
                     /note="hypoxanthine-guanine phosphoribosyltransferase"
BASE COUNT        9 a     10 c      7 g     16 t
ORIGIN

  Query Match             67.0%;  Score 13.4;  DB 9;  Length 42;
  Best Local Similarity   93.3%;  Pred. No. 3.5e+04;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      2 TGCTCCCTGTTTCAT 16
          ||| ||||||||||
Db     12 TGCACCCTGTTTCAT 26


RESULT 3
AR111310
LOCUS       AR111310                  93 bp    DNA     linear   PAT 14-FEB-2001
DEFINITION  Sequence 425 from patent US 6127119.
ACCESSION   AR111310
VERSION     AR111310.1  GI:12828158
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 93)
  AUTHORS   Stephens,A., Gold,L. and Speck,U.
  TITLE     Nucleic acid ligands of tissue target
  JOURNAL   Patent: US 6127119-A 425 03-OCT-2000;
FEATURES             Location/Qualifiers
     source          1. .93
                     /organism="unknown"
BASE COUNT       29 a     26 c     20 g     18 t
ORIGIN

  Query Match             67.0%;  Score 13.4;  DB 6;  Length 93;
  Best Local Similarity   93.3%;  Pred. No. 3.4e+04;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      6 CCCTGTTTCATCAAT 20
          ||||||||||| ||||
Db     38 CCCTGTTTCAACAAT 52


RESULT 4
AR123848/c
LOCUS       AR123848                  49 bp    DNA     linear   PAT 16-MAY-2001
DEFINITION  Sequence 26 from patent US 6171816.
ACCESSION   AR123848
VERSION     AR123848.1  GI:14109209
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 49)
  AUTHORS   Yu,G.-L., Dillon,P.J., Ebner,R. and Endress,G.A.
  TITLE     Human XAG-1 polynucleotides and polypeptides
  JOURNAL   Patent: US 6171816-A 26 09-JAN-2001;
FEATURES             Location/Qualifiers
     source          1. .49
                     /organism="unknown"
BASE COUNT       15 a     11 c     10 g     13 t
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 6;  Length 49;
```

```
  Best Local Similarity   83.3%;  Pred. No. 4.5e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 GTGCTCCCTGTTTCATCA 18
          |||| || | ||||||||
Db     33 GTGCCCCTTCTTTCATCA 16


RESULT 5
AX103651/c
LOCUS       AX103651                  57 bp    DNA     linear   PAT 30-APR-2001
DEFINITION  Sequence 15 from Patent WO0121662.
ACCESSION   AX103651
VERSION     AX103651.1  GI:13919874
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 57)
  AUTHORS   Habermann,P. and Bender,R.
  TITLE     Signal sequences for the production of leu-hirudine via secretion
            by E. coli in a culture medium
  JOURNAL   Patent: WO 0121662-A 15 29-MAR-2001;
            Aventis Pharma Deutschland GmbH (DE)
FEATURES             Location/Qualifiers
     source          1. .57
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="misc feature"
BASE COUNT       11 a     14 c     15 g     17 t
ORIGIN

  Query Match            66.0%;  Score 13.2;  DB 6;  Length 57;
  Best Local Similarity   83.3%;  Pred. No. 4.5e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      3 GCTCCCTGTTTCATCAAT 20
          || | |||||||||| |||
Db     26 GCGCTCTGTTTCATGAAT 9


RESULT 6
ECO5ERNAC/c
LOCUS       ECO5ERNAC                 61 bp    DNA     linear   BCT 26-APR-1993
DEFINITION  E.coli 5'-termini mRNA of the ileS-lsp genes, SM31-2B4 a/b.
ACCESSION   M16642
VERSION     M16642.1  GI:145142
KEYWORDS    .
SOURCE      E.coli (strain CS412), cDNA to mRNA, clone SM31-2B4 a/b.
  ORGANISM  Escherichia coli
            Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
            Escherichia.
REFERENCE   1  (bases 1 to 61)
  AUTHORS   Miller,K.W., Bouvier,J., Stragier,P. and Wu,H.C.
  TITLE     Identification of the genes in the Escherichia coli ileS-lsp
            operon. Analysis of multiple polycistronic mRNAs made in vivo
  JOURNAL   J. Biol. Chem. 262 (15), 7391-7397 (1987)
  MEDLINE   87222348
   PUBMED   3294831
FEATURES             Location/Qualifiers
     source          1. .61
                     /organism="Escherichia coli"
                     /db_xref="taxon:562"
     mRNA            1. .>61
                     /product="ileS-lsp mRNA"
BASE COUNT       11 a     15 c     18 g     17 t
ORIGIN      1 min on K12 map.

  Query Match            66.0%;  Score 13.2;  DB 1;  Length 61;
  Best Local Similarity   83.3%;  Pred. No. 4.5e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 GTGCTCCCTGTTTCATCA 18
          |||| |||| |||||||
Db     47 GTGCGCCCTAATTCATCA 30


RESULT 7
AX402155
LOCUS       AX402155                  21 bp    DNA     linear   PAT 07-JUN-2002
DEFINITION  Sequence 17 from Patent WO0226813.
ACCESSION   AX402155
VERSION     AX402155.1  GI:21387298
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Mao,Y. and Xie,Y.
  TITLE     A novel peptide - human muscle cell enhanced associative factor
            222.88 and the polynucleotide coding this novel peptide
  JOURNAL   Patent: WO 0226813-A 17 04-APR-2002;
            SHANGHAI BIOWINDOW GENE DEV IN (CN)
FEATURES             Location/Qualifiers
     source          1. .21
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT        4 a      6 c      4 g      7 t
ORIGIN

  Query Match            65.0%;  Score 13;  DB 6;  Length 21;
  Best Local Similarity   100.0%;  Pred. No. 6e+04;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      7 CCTGTTTCATCAA 19
          |||||||||||||
Db      2 CCTGTTTCATCAA 14


RESULT 8
AR039988/c
LOCUS       AR039988                  27 bp    DNA     linear   PAT 29-SEP-1999
DEFINITION  Sequence 836 from patent US 5807743.
ACCESSION   AR039988
VERSION     AR039988.1  GI:5959351
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 27)
  AUTHORS   Stinchcomb,D.T. and McSwiggen,J.A.
  TITLE     Interleukin-2 receptor gamma-chain ribozymes
  JOURNAL   Patent: US 5807743-A 836 15-SEP-1998;
FEATURES             Location/Qualifiers
     source          1. .27
                     /organism="unknown"
BASE COUNT        8 a      4 c      9 g      5 t      1 others
ORIGIN

  Query Match            65.0%;  Score 13;  DB 6;  Length 27;
  Best Local Similarity   92.9%;  Pred. No. 5.9e+04;
  Matches   13;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      5 TCCCTGTTTCATCA 18
          |||||||||| |||
Db     26 TCCCTGTTTCNTCA 13


RESULT 9
AR189295/c
LOCUS       AR189295                  27 bp    DNA     linear   PAT 20-APR-2002
DEFINITION  Sequence 4783 from patent US 6346398.
ACCESSION   AR189295
VERSION     AR189295.1  GI:20235260
KEYWORDS    .
```

```
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 27)
  AUTHORS   Pavco,P., McSwiggen,J., Stinchcomb,D. and Escobedo,J.
  TITLE     Method and reagent for the treatment of diseases or conditions
            related to levels of vascular endothelial growth factor receptor
  JOURNAL   Patent: US 6346398-A 4783 12-FEB-2002;
FEATURES             Location/Qualifiers
     source          1. .27
                     /organism="unknown"
BASE COUNT        9 a      6 c      8 g      3 t      1 others
ORIGIN

  Query Match            65.0%;  Score 13;  DB 6;  Length 27;
  Best Local Similarity   92.9%;  Pred. No. 5.9e+04;
  Matches   13;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       5 TCCCTGTTTCATCA 18
           |||||||||| |||
Db      26 TCCCTGTTTCNTCA 13


RESULT 10
AR070500/c
LOCUS       AR070500                  30 bp    DNA     linear   PAT 18-FEB-2000
DEFINITION  Sequence 14 from patent US 5906819.
ACCESSION   AR070500
VERSION     AR070500.1  GI:7221388
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 30)
  AUTHORS   Kaibuchi,K., Iwamatsu,A., Nakano,T., Ito,M. and Takahashi,N.
  TITLE     Rho target protein Rho-kinase
  JOURNAL   Patent: US 5906819-A 14 25-MAY-1999;
FEATURES             Location/Qualifiers
     source          1. .30
                     /organism="unknown"
BASE COUNT       10 a      3 c     11 g      6 t
ORIGIN

  Query Match            64.0%;  Score 12.8;  DB 6;  Length 30;
  Best Local Similarity   87.5%;  Pred. No. 7.6e+04;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       5 TCCCTGTTTCATCAAT 20
           | ||||||||||| ||
Db      27 TGCCTGTTTCATCCAT 12


RESULT 11
AX128522/c
LOCUS       AX128522                  30 bp    DNA     linear   PAT 15-MAY-2001
DEFINITION  Sequence 30 from Patent WO0131014.
ACCESSION   AX128522
VERSION     AX128522.1  GI:14135003
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 30)
  AUTHORS   Vogeli,G., Wood,L.S. and Merchant,K.
  TITLE     G protein-coupled receptors expressed in brain
  JOURNAL   Patent: WO 0131014-A 30 03-MAY-2001;
            PHARMACIA & UPJOHN COMPANY (US)
FEATURES             Location/Qualifiers
     source          1. .30
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="Primer LW1405"
BASE COUNT       14 a      4 c      7 g      5 t
ORIGIN

  Query Match            64.0%;  Score 12.8;  DB 6;  Length 30;
  Best Local Similarity   87.5%;  Pred. No. 7.6e+04;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       5 TCCCTGTTTCATCAAT 20
           | ||||||||||| ||
Db      25 TTCCTGTTTCATCCAT 10


RESULT 12
E26044/c
LOCUS       E26044                    30 bp    DNA     linear   PAT 18-JUN-2001
DEFINITION  Peptide fragment of neutralizing antibody against serine protease
            derived from hepatitis C virus.
ACCESSION   E26044
VERSION     E26044.1  GI:13025025
KEYWORDS    JP 1999127861-A/6.
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 30)
  AUTHORS   Satoru,M., Yoichi,O. and Takamasa,U.
  TITLE     Peptide fragment of neutralizing antibody against serine protease
            derived from hepatitis C virus
  JOURNAL   Patent: JP 1999127861-A 6 18-MAY-1999;
            JAPAN ENERGY CORP
COMMENT     OS   Unidentified
            PN   JP 1999127861-A/6
            PD   18-MAY-1999
            PF   29-OCT-1997 JP 1997297451
            PR
            PI   SATORU MISAWA,YOICHI OBA,TAKAMASA UENO
            PC   C12N15/09,C07K7/06,C07K7/08,C07K16/38,C12N5/10, PC
            C12N9/99//A61K38/55,
            PC   A61K39/395,(C12N15/09,C12R1:92),C12N15/00,C12N5/00,A61K37/64,
            PC   (C12N15/00,C12R1:92)
            CC   Strandedness: Single;
            CC   Topology: Linear;
            FH   Key             Location/Qualifiers
            FT   source          1. .30
            FT                   /organism='Unidentified'.
FEATURES             Location/Qualifiers
     source          1. .30
                     /organism="unidentified"
                     /db_xref="taxon:32644"
BASE COUNT        6 a      10 c      7 g      6 t      1 others
ORIGIN

  Query Match            64.0%;  Score 12.8;  DB 6;  Length 30;
  Best Local Similarity   77.8%;  Pred. No. 7.6e+04;
  Matches   14;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 GTGCTCCCTGTTTCATCA 18
           ||||  |||| |||||||
Db      25 GTGCAGCCTGKGTCATCA 8


RESULT 13
AR195964/c
LOCUS       AR195964                  31 bp    DNA     linear   PAT 20-APR-2002
DEFINITION  Sequence 429 from patent US 6350934.
ACCESSION   AR195964
VERSION     AR195964.1  GI:20245401
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 31)
  AUTHORS   Zwick,M.G., Edington,B.E., McSwiggen,J.A., Merlo,P.Ann.Owens.,
            Guo,L., Skokut,T.A., Young,S.A., Folkerts,O. and Merlo,D.J.
  TITLE     Nucleic acid encoding delta-9 desaturase
```

```
  JOURNAL   Patent: US 6350934-A 429 26-FEB-2002;
FEATURES             Location/Qualifiers
     source          1. .31
                     /organism="unknown"
BASE COUNT        8 a      6 c     12 g      4 t      1 others
ORIGIN

  Query Match             64.0%;  Score 12.8;  DB 6;  Length 31;
  Best Local Similarity   82.4%;  Pred. No. 7.6e+04;
  Matches   14;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 TGCTCCCTGTTTCATCA 18
          ||||| | ||||| |||
Db     31 TGCTCGCCGTTTCNTCA 15


RESULT 14
AX108135/c
LOCUS       AX108135                 43 bp    DNA     linear   PAT 30-APR-2001
DEFINITION  Sequence 129 from Patent WO0125267.
ACCESSION   AX108135
VERSION     AX108135.1  GI:13923461
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 43)
  AUTHORS   Braun,C., Purac,A. and Borgford,T.
  TITLE     Improved ricin-like toxins for treatment of cancer
  JOURNAL   Patent: WO 0125267-A 129 12-APR-2001;
            Twinstrand Therapeutics, Inc. (CA)
FEATURES             Location/Qualifiers
     source          1. .43
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="synthetic construct"
BASE COUNT       16 a      5 c     14 g      8 t
ORIGIN

  Query Match             64.0%;  Score 12.8;  DB 6;  Length 43;
  Best Local Similarity   87.5%;  Pred. No. 7.4e+04;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      2 TGCTCCCTGTTTCATC 17
          | ||||| ||||||||
Db     29 TCCTCCCGGTTTCATC 14


RESULT 15
AF274789/c
LOCUS       AF274789                 84 bp    DNA     linear   PRI 04-JUL-2000
DEFINITION  Homo sapiens clone 1105 T cell receptor beta chain (TCRB) gene,
            partial cds.
ACCESSION   AF274789
VERSION     AF274789.1  GI:8926458
KEYWORDS    .
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (sites)
  AUTHORS   Dulphy,N., Peyrat,M.-A., Tieng,V., Douay,C., Rabian,C., Tamouza,R.,
            Laoussadi,S., Berenbaum,F., Chabot,A., Bonneville,M., Charron,D.
            and Toubert,A.
  TITLE     Common intra-articular T cell expansions in patients with reactive
            arthritis: identical beta-chain junctional sequences and
            cytotoxicity toward HLA-B27
  JOURNAL   J. Immunol. 162 (7), 3830-3839 (1999)
  MEDLINE   99218398
   PUBMED   10201900
REFERENCE   2  (bases 1 to 84)
  AUTHORS   May,E.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-JUN-2000) Internal Med/Rheumatology, University of
            Texas Southwestern Medical School, 5323 Harry Hines Blvd, Dallas,
            TX 75235-8884, USA
FEATURES             Location/Qualifiers
     source          1. .84
                     /organism="Homo sapiens"
                     /isolate="patient CN"
                     /db_xref="taxon:9606"
                     /clone="1105"
                     /tissue_type="synovial fluid"
                     /note="isolated from a HLA-B27-positive patient with
                     Reactive Arthritis"
     gene            <1. .>84
                     /gene="TCRB"
                     /note="TCRBV11J2S3"
     mRNA            <1. .>84
                     /gene="TCRB"
                     /product="T cell receptor beta chain"
     CDS             <1. .>84
                     /gene="TCRB"
                     /codon_start=3
                     /product="T cell receptor beta chain"
                     /protein_id="AAF81838.1"
                     /db_xref="GI:8926459"
                     /translation="ARPSHTSQYLCASNETGGTDTQYFGPA"
BASE COUNT       20 a     26 c     21 g     17 t
ORIGIN

  Query Match             64.0%;  Score 12.8;  DB 9;  Length 84;
  Best Local Similarity   87.5%;  Pred. No. 7.2e+04;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 GTGCTCCCTGTTTCAT 16
          |||| |||||| ||||
Db     58 GTGCCCCCTGTCTCAT 43


Search completed: December  3, 2002, 18:14:00
Job time : 309.2 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 05:52:31 ; Search time 98.55 Seconds
(without alignments)
457.027 Million cell updates/sec

Title: US-09-296-264-18
Perfect score: 20
Sequence: 1 gtgctccctgtttcatcaat 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 2390332

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	20	100.0	20	21	AAZ31448	Human neuropilin m
	2	14.4	72.0	27	19	AAV38846	HCV core gene anti
c	3	14.2	71.0	25	21	AAA07857	Forward RH mapping
c	4	14.2	71.0	25	21	AAA07879	Primer for Hs-UNC-
	5	13.6	68.0	57	24	ABK49548	Oligonucleotide fo
c	6	13.6	68.0	94	22	ABA69960	Human foetal liver
c	7	13.6	68.0	94	22	ABA36797	Probe #15263 for g
c	8	13.6	68.0	94	22	AAK18170	Human brain expres
c	9	13.6	68.0	94	22	AAK44062	Human bone marrow
c	10	13.6	68.0	94	22	AAI24710	Probe #14643 for g
c	11	13.6	68.0	94	22	AAI50072	Probe #18758 used
c	12	13.6	68.0	94	24	ABS18295	Human genome-deriv
	13	13.6	68.0	99	23	AAS50985	Staphylococcus aur
c	14	13.4	67.0	60	24	ABN46831	Human spliced tran
	15	13.4	67.0	93	20	AAX85482	Human artery perfu
c	16	13.2	66.0	24	24	ABK87125	Human KIAA 96 mark
c	17	13.2	66.0	33	24	ABL53499	Human cyclin G29.1
c	18	13.2	66.0	47	21	AAZ68040	Human map-related
c	19	13.2	66.0	49	19	AAV19177	Vector pQE60-speci
c	20	13.2	66.0	49	22	AAF63322	PCR primer specifi
c	21	13.2	66.0	57	22	AAF61500	Leu-hirudin/alkali
c	22	13.2	66.0	60	24	ABN43456	Human spliced tran
	23	13.2	66.0	60	24	ABN58830	Human spliced tran
	24	13	65.0	21	24	ABK92279	Human tumour suppr
c	25	13	65.0	27	18	AAX72033	Mouse flk-1 VEGF r
c	26	13	65.0	27	19	AAV94303	Mouse IL-2 recepto
	27	13	65.0	47	21	AAZ66137	Human map-related
	28	12.8	64.0	19	21	AAZ76078	Human biallelic ma
	29	12.8	64.0	20	21	AAA60151	Human PPARalpha ge
	30	12.8	64.0	20	24	ABL45403	Human chromosome 2
	31	12.8	64.0	29	20	AAX59362	Human secreted pro
	32	12.8	64.0	29	22	AAS59330	Human secreted pro
	33	12.8	64.0	29	24	ABA90999	Biotinylated oligo
c	34	12.8	64.0	30	20	AAX57789	Coding region for
c	35	12.8	64.0	30	22	AAD06518	Human CON166 G pro
c	36	12.8	64.0	31	18	AAX62554	Granule bound star
	37	12.8	64.0	31	20	AAX38847	Human genomic DNA
c	38	12.8	64.0	33	18	AAT97920	Upstream primer ri
c	39	12.8	64.0	43	20	AAX04249	Proricin productio
c	40	12.8	64.0	43	22	AAH77625	Castor bean plant
c	41	12.8	64.0	47	21	AAZ69450	Human map-related
	42	12.8	64.0	60	18	AAV76470	Staphylococcus aur
c	43	12.8	64.0	60	24	ABN44200	Human spliced tran
	44	12.8	64.0	60	24	ABN44791	Human spliced tran
	45	12.8	64.0	65	24	ABN55940	Mouse spliced tran

ALIGNMENTS

RESULT 1
AAZ31448
ID AAZ31448 standard; DNA; 20 BP.
XX
AC AAZ31448;
XX
DT 07-FEB-2000 (first entry)
XX
DE Human neuropilin mRNA specific antisense oligo GTI3619.
XX
KW Neuropilin; human; growth; metastasis; tumor; neovascularisation;
KW cancer; papilloma; diabetic retinopathy; antisense; ss.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9955855-A2.
XX
PD 04-NOV-1999.
XX
PF 23-APR-1999; 99WO-CA00324.
XX
PR 23-APR-1998; 98US-0082791.
XX
PA (GENE-) GENESENSE TECHNOLOGIES INC.
XX
PI Wright JA, Young AH, Lee YS;
XX
DR WPI; 2000-023357/02.
XX
PT Antisense oligonucleotides that inhibit neuropilin expression, useful
PT for treating cancer -

XX
PS Claim 4; Page 16; 57pp; English.
XX
CC Sequences AAZ31431-460 represent antisense oligonucleotides which
CC inhibit human neuropilin expression. The antisense oligonucleotides can
CC be used to inhibit the growth or metastasis of a mammalian tumor and
CC inhibit neovascularisation. The oligonucleotides may be used to treat
CC various forms of cancers or tumors, such as sarcomas, melanomas,
CC adenomas, carcinomas of solid tissue, hypoxic tumors, squamous cell
CC carcinomas of the mouth, throat, larynx and lung, genitourinary cancers
CC such as cervical and bladder cancer, hematopoietic cancers, colon cancer,
CC breast cancer, pancreatic cancer, renal cancer, brain cancer, skin
CC cancer, liver cancer, head and neck cancers, and nervous system cancers,
CC as well as benign lesions such as papillomas. The methods may be used to
CC treat neovascularisation disorders such as diabetic retinopathy, and
CC retinopathy of prematurity and age related macular degeneration.
XX
SQ Sequence 20 BP; 3 A; 6 C; 3 G; 8 T; 0 other;

Query Match 100.0%; Score 20; DB 21; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.8;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTGCTCCCTGTTTCATCAAT 20
||||||||||||||||||||
Db 1 GTGCTCCCTGTTTCATCAAT 20

RESULT 2
AAV38846
ID AAV38846 standard; DNA; 27 BP.
XX
AC AAV38846;
XX
DT 26-OCT-1998 (first entry)
XX
DE HCV core gene antisense PCR primer #155.
XX
KW Hepatitis delta virus; L-HDAg; virus-like particle; infection; HCV;
KW core protein; hepatitis B virus surface antigen; HBsAg; immunogen;
KW vaccine; PCR; primer; ss.
XX
OS Synthetic.
OS Hepatitis C virus.
XX
PN WO9828004-A1.
XX
PD 02-JUL-1998.
XX
PF 24-DEC-1997; 97WO-AU00884.
XX
PR 24-DEC-1996; 96AU-0004341.
XX
PA (QUEE-) QUEENSLAND DEPT HEALTH SAKZEWSKI VIRUS.
XX
PI Gowans EJ, MacNaughton TB;
XX
DR WPI; 1998-377411/32.
XX
PT Virus-like particle for, e.g. treating microbial infection -
PT comprises polypeptide from microorganism and sequence from Hepatitis
PT D virus large protein, partially enveloped by Hepatitis B surface
PT antigen
XX
PS Example; Page 18; 72pp; English.
XX
CC Antisense primer #155 comprises 7 nucleotides designed to
CC facilitate cloning followed by nucleotides 501-482 of the
CC hepatitis C virus (HCV) genome (nucleotide +1 representing the
CC start of the long open reading frame). It was used with sense
CC primer #156 (see AAV38845) in a PCR amplification of the HCV core
CC gene using products of a SISPA reaction (see AAV38843-44) as
CC template. The 531-nucleotide product of the PCR was cloned into
CC pBluescript KS to produce clone pA2. The sequence of the HCV cDNA
CC insert in pA2 is provided in AAV42303. The HCV core protein can be
CC used as an immunogen in novel fusion proteins (see AAW62657-59) that
CC comprise HCV core protein and at least 19 amino acids (see AAW62827)
CC of the C-terminal sequence of the large protein from hepatitis D
CC virus (L-HDAg). In novel virus-like particles of the invention, a
CC fusion protein immunogen is at least partially enveloped by
CC hepatitis B surface antigen. The virus-like particle is used to
CC ameliorate or protect against infections caused by hepatitis B
CC virus and/or another microorganism, especially HCV.
XX
SQ Sequence 27 BP; 3 A; 9 C; 7 G; 8 T; 0 other;

Query Match 72.0%; Score 14.4; DB 19; Length 27;
Best Local Similarity 93.8%; Pred. No. 1.1e+03;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GTGCTCCCTGTTTCAT 16
|||||||||||| |||
Db 4 GTGCTCCCTGTTGCAT 19

RESULT 3
AAA07857/c
ID AAA07857 standard; DNA; 25 BP.
XX
AC AAA07857;
XX
DT 07-JUL-2000 (first entry)
XX
DE Forward RH mapping primer for Hs-UNC-53/3.
XX
KW UNC-53; Caenorhabditis elegans; microtubule; neural regeneration;
KW anticancer; anti-neurodegeneration; antifibrotic; anti-adhesive; human;
KW antisclerotic; antimetastatic; anti-arthritic; autoimmune disease;
KW PCR primer; radiation hybrid mapping; ss.
XX
OS Homo sapiens.
XX
PN WO9963080-A1.
XX
PD 09-DEC-1999.
XX
PF 02-JUN-1999; 99WO-EP03848.
XX
PR 03-JUN-1998; 98GB-0011962.
XX
PA (JANC) JANSSEN PHARM NV.
XX
PI Luyten WHML, De Raeymaeker MC, Geysen JJGH, Bogaert TAOE;
PI Maerten LJS, Verhasselt P, Van De Craen M;
XX
DR WPI; 2000-116370/10.
XX
PT Novel proteins and nucleic acids e.g. for treating neurodegeneration -
XX
PS Disclosure; Page 47; 146pp; English.
XX
CC The invention provides vertebrate (human) protein homologue of a UNC-53
CC protein of Caenorhabditis elegans. The UNC-53 binds to microtubules or
CC their plus ends. The UNC-53 sequences are used to promote neural
CC regeneration, revascularization and wound healing; also for treating
CC neurodegenerative disease, acute traumatic injury, fibrotic disease and
CC autoimmune diseases (e.g. rheumatoid arthritis and sclerosis). The
CC UNC-53 polynucleotides can be used for recombinant production of the
CC proteins, as a source of probes for detecting allelic variants and
CC polymorphisms, for sequencing genomic DNA and for detecting UNC-53
CC expression; and as source of therapeutic antisense sequences. Cells that
CC express the protein are used to identify regulators of cell shape,
CC growth, motility and migration. They can also be used to identify
CC proteins that are involved in signal transduction pathways also involving
CC UNC-53, and to identify compounds that alter attachment of UNC-53 to
CC microtubules. A target gene coupled to a UNC-53 encoding sequence may be

```
CC   used to deliver the target gene to a cellular microtubule or its plus
CC   ends. Sequences AAA07851-7866 represent PCR primers used in radiation
CC   hybrid (RH) mapping and FISH analysis for chromosomal localisation of
CC   human UNC-53.
XX
SQ   Sequence 25 BP; 8 A; 6 C; 6 G; 5 T; 0 other;

  Query Match             71.0%;  Score 14.2;  DB 21;  Length 25;
  Best Local Similarity   84.2%;  Pred. No. 1.4e+03;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 GTGCTCCCTGTTTCATCAA 19
          | |||| |||||||| |||
Db     21 GAGCTCTCTGTTTCAGCAA 3


RESULT 4
AAA07879/c
ID   AAA07879 standard; DNA; 25 BP.
XX
AC   AAA07879;
XX
DT   07-JUL-2000  (first entry)
XX
DE   Primer for Hs-UNC-53/3 variant validation.
XX
KW   UNC-53; Caenorhabditis elegans; microtubule; neural regeneration;
KW   anticancer; anti-neurodegeneration; antifibrotic; anti-adhesive; human;
KW   antisclerotic; antimetastatic; anti-arthritic; autoimmune disease;
KW   PCR primer; ss.
XX
OS   Homo sapiens.
XX
PN   WO9963080-A1.
XX
PD   09-DEC-1999.
XX
PF   02-JUN-1999;   99WO-EP03848.
XX
PR   03-JUN-1998;   98GB-0011962.
XX
PA   (JANC ) JANSSEN PHARM NV.
XX
PI   Luyten WHML,  De Raeymaeker MC,  Geysen JJGH,  Bogaert TAOE;
PI   Maerten LJS,  Verhasselt P,  Van De Craen M;
XX
DR   WPI; 2000-116370/10.
XX
PT   Novel proteins and nucleic acids e.g. for treating neurodegeneration  -
XX
PS   Disclosure; Page 51; 146pp; English.
XX
CC   The invention provides vertebrate (human) protein homologue of a UNC-53
CC   protein of Caenorhabditis elegans. The UNC-53 binds to microtubules or
CC   their plus ends. The UNC-53 sequences are used to promote neural
CC   regeneration, revascularization and wound healing; also for treating
CC   neurodegenerative disease, acute traumatic injury, fibrotic disease and
CC   autoimmune diseases (e.g. rheumatoid arthritis and sclerosis). The
CC   UNC-53 polynucleotides can be used for recombinant production of the
CC   proteins, as a source of probes for detecting allelic variants and
CC   polymorphisms, for sequencing genomic DNA and for detecting UNC-53
CC   expression; and as source of therapeutic antisense sequences. Cells that
CC   express the protein are used to identify regulators of cell shape,
CC   growth, motility and migration. They can also be used to identify
CC   proteins that are involved in signal transduction pathways also involving
CC   UNC-53, and to identify compounds that alter attachment of UNC-53 to
CC   microtubules. A target gene coupled to a UNC-53 encoding sequence may be
CC   used to deliver the target gene to a cellular microtubule or its plus
CC   ends. Sequences AAA07877-7880 represent PCR primers used for validating
CC   the variants at the 5' end of the Hs-UNC-53/3 cDNA sequence.
XX
SQ   Sequence 25 BP; 8 A; 6 C; 6 G; 5 T; 0 other;

  Query Match             71.0%;  Score 14.2;  DB 21;  Length 25;
  Best Local Similarity   84.2%;  Pred. No. 1.4e+03;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 GTGCTCCCTGTTTCATCAA 19
          | |||| |||||||| |||
Db     21 GAGCTCTCTGTTTCAGCAA 3


RESULT 5
ABK49548
ID   ABK49548 standard; DNA; 57 BP.
XX
AC   ABK49548;
XX
DT   15-JUL-2002  (first entry)
XX
DE   Oligonucleotide for partial ovalbumin gene, OVA3R.
XX
KW   Ovalbumin; immune tolerance; transgenic plant;
KW   T-cell epitope; allergic protein; ss.
XX
OS   Synthetic.
XX
PN   JP2002085068-A.
XX
PD   26-MAR-2002.
XX
PF   08-SEP-2000; 2000JP-0272863.
XX
PR   08-SEP-2000; 2000JP-0272863.
XX
PA   (SUMO ) SUMITOMO CHEM CO LTD.
XX
DR   WPI; 2002-378273/41.
XX
PT   Induction of immune tolerance by oral administration of an effective
PT   amount of transformed plant comprising a DNA coding for a polypeptide
PT   having an amino acid sequence of T-cell epitope of an allergic protein
PT   -
XX
PS   Example 2; Page 15; 19pp; Japanese.
XX
CC   The invention describes a method for induction of immune tolerance
CC   in animals, including humans. The treatment involves by oral
CC   administration of an effective amount of transformed plant comprising
CC   a DNA coding for a polypeptide having an amino acid sequence of T-cell
CC   epitope of an allergic protein. The introduced protein is not degraded
CC   by digestive enzymes in the host. This oligonucleotide sequence is
CC   used in an expression vector for plant transformation and encodes a
CC   fragment of the ovalbumin gene.
XX
SQ   Sequence 57 BP; 14 A; 12 C; 11 G; 20 T; 0 other;

  Query Match             68.0%;  Score 13.6;  DB 24;  Length 57;
  Best Local Similarity   80.0%;  Pred. No. 3e+03;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 GTGCTCCCTGTTTCATCAAT 20
          || || |||| ||||| |||
Db      4 GTTCTACCTGCTTCATTAAT 23


RESULT 6
ABA69960/c
ID   ABA69960 standard; DNA; 94 BP.
XX
AC   ABA69960;
XX
DT   01-FEB-2002  (first entry)
XX
DE   Human foetal liver single exon nucleic acid probe #18265.
XX
```

KW Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
XX
OS Homo sapiens.
XX
PN WO200157277-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00669.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-483447/52.
XX
PT Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human fetal liver -
XX
PS Claim 4; SEQ ID NO 18265; 639pp + sequence listing; English.
XX
CC The invention relates to a single exon nucleic acid probe for
CC measuring human gene expression in a sample derived from human foetal
CC liver. The single exon nucleic acid probes may be used for predicting,
CC measuring and displaying gene expression in samples derived from human
CC fetal liver. The present sequence is a single exon nucleic acid
CC probe of the invention.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 94 BP; 24 A; 11 C; 22 G; 37 T; 0 other;

Query Match 68.0%; Score 13.6; DB 22; Length 94;
Best Local Similarity 80.0%; Pred. No. 3.2e+03;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 GTGCTCCCTGTTTCATCAAT 20
||||||||||||||||||||
Db 29 GTGCTCACTTTTTCTTAAAT 10

RESULT 7
ABA36797/c
ID ABA36797 standard; DNA; 94 BP.
XX
AC ABA36797;
XX
DT 23-JAN-2002 (first entry)
XX
DE Probe #15263 for gene expression analysis in human heart cell sample.
XX
KW Human; gene expression; heart; microarray; vascular system; probe;
KW cardiovascular disease; hypertension; cardiac arrhythmia;
KW congenital heart disease; ss.
XX
OS Homo sapiens.
XX
PN WO200157274-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00666.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488899/53.
XX
PT Single exon nucleic acid probes for analyzing gene expression in human
PT hearts -
XX
PS Claim 4; SEQ ID No 15263; 530pp; English.
XX
CC The present invention relates to single exon nucleic acid probes for
CC measuring human gene expression in a sample derived from human heart. The
CC present sequence is one such probe. The probes may be used for
CC predicting, measuring and displaying gene expression in samples derived
CC from the human heart via microarrays. By measuring gene expression, the
CC probes are useful for predicting, diagnosing, grading, staging,
CC monitoring and prognosing diseases of the human heart and vascular system
CC e.g. cardiovascular disease, hypertension, cardiac arrhythmias and
CC congenital heart disease.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 94 BP; 24 A; 11 C; 22 G; 37 T; 0 other;

Query Match 68.0%; Score 13.6; DB 22; Length 94;
Best Local Similarity 80.0%; Pred. No. 3.2e+03;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 GTGCTCCCTGTTTCATCAAT 20
||||||||||||||||||||
Db 29 GTGCTCACTTTTTCTTAAAT 10

RESULT 8
AAK18170/c
ID AAK18170 standard; DNA; 94 BP.
XX
AC AAK18170;
XX
DT 05-NOV-2001 (first entry)
XX
DE Human brain expressed single exon probe SEQ ID NO: 18161.
XX
KW Human; brain expressed exon; gene expression analysis; probe;
KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
KW epilepsy; cancer; ss.
XX
OS Homo sapiens.
XX
PN WO200157275-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00667.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.

```
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-483446/52.
XX
PT   Single exon nucleic acid probes for analyzing gene expression in human
PT   brains -
XX
PS   Example 4; SEQ ID NO: 18161; 650pp + Sequence Listing; English.
XX
CC   The present invention provides a number of single exon nucleic acid
CC   probes which are derived from genomic sequences expressed in the human
CC   brain. They can be used to measure gene expression in brain cell samples,
CC   which may enable the diagnosis and improved treatment of nervous system
CC   diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC   epilepsy and cancers. The present sequence is one of the probes of the
CC   invention.
XX
SQ   Sequence 94 BP; 24 A; 11 C; 22 G; 37 T; 0 other;

  Query Match             68.0%;  Score 13.6;  DB 22;  Length 94;
  Best Local Similarity   80.0%;  Pred. No. 3.2e+03;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 GTGCTCCCTGTTTCATCAAT 20
           |||||| || |||| | |||
Db      29 GTGCTCACTTTTTCTTAAAT 10


RESULT 9
AAK44062/c
ID   AAK44062 standard; DNA; 94 BP.
XX
AC   AAK44062;
XX
DT   06-NOV-2001  (first entry)
XX
DE   Human bone marrow expressed single exon probe SEQ ID NO: 18619.
XX
KW   Human; bone marrow expressed exon; gene expression analysis; probe;
KW   microarray; cancer; leukaemia; lymphoma; myeloma; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157276-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00668.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488900/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for
PT   analyzing gene expression in human bone marrow -
XX
PS   Example 4; SEQ ID NO: 18619; 658pp + Sequence Listing; English.
XX
CC   The present invention provides a number of single exon nucleic acid
CC   probes which are derived from genomic sequences expressed in the human
CC   bone marrow. They can be used to measure gene expression in bone marrow
CC   samples, which may enable the improved diagnosis and treatment of cancers
CC   such as lymphoma, leukaemia and myeloma. The present sequence is one of
CC   the probes of the invention.
XX
SQ   Sequence 94 BP; 24 A; 11 C; 22 G; 37 T; 0 other;

  Query Match             68.0%;  Score 13.6;  DB 22;  Length 94;
  Best Local Similarity   80.0%;  Pred. No. 3.2e+03;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 GTGCTCCCTGTTTCATCAAT 20
           |||||| || |||| | |||
Db      29 GTGCTCACTTTTTCTTAAAT 10


RESULT 10
AAI24710/c
ID   AAI24710 standard; DNA; 94 BP.
XX
AC   AAI24710;
XX
DT   12-OCT-2001  (first entry)
XX
DE   Probe #14643 for gene expression analysis in human cervical cell sample.
XX
KW   Probe; human; microarray; gene expression; cervical epithelial cell;
KW   cervical cancer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157278-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00670.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488901/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for
PT   analyzing gene expression in human cervical epithelial cells -
XX
PS   Claim 25; SEQ ID No 14643; 487pp; English.
XX
CC   The present invention relates to human single exon nucleic acid probes
CC   (SENP). The present sequence is one such probe. The SENPs are derived
CC   from human HeLa cells. The SENPs can be used to produce a single exon
CC   microarray, which can be used for measuring human gene expression in a
CC   sample derived from human cervical epithelial cells. By measuring gene
CC   expression, the probes are therefore useful in grading and/or staging
CC   of diseases of the cervix, notably cervical cancer.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 94 BP; 24 A; 11 C; 22 G; 37 T; 0 other;

  Query Match             68.0%;  Score 13.6;  DB 22;  Length 94;
  Best Local Similarity   80.0%;  Pred. No. 3.2e+03;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 GTGCTCCCTGTTTCATCAAT 20
           |||||| || |||| | |||
```

```
Db     29 GTGCTCACTTTTTCTTAAAT 10


RESULT 11
AAI50072/c
ID   AAI50072 standard; DNA; 94 BP.
XX
AC   AAI50072;
XX
DT   17-OCT-2001  (first entry)
XX
DE   Probe #18758 used to measure gene expression in human placenta sample.
XX
KW   Probe; microarray; human; placenta; antenatal diagnosis;
KW   genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157272-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00663.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488897/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for
PT   analyzing gene expression in human placenta -
XX
PS   Claim 25; SEQ ID No 18758; 654pp; English.
XX
CC   The present invention relates to single exon nucleic acid probes (SENP).
CC   The present sequence is one such probe. The probes are useful for
CC   producing a microarray for predicting, measuring and displaying gene
CC   expression in samples derived from human placenta. The probes are useful
CC   for antenatal diagnosis of human genetic disorders.
XX
SQ   Sequence 94 BP; 24 A; 11 C; 22 G; 37 T; 0 other;

  Query Match             68.0%;  Score 13.6;  DB 22;  Length 94;
  Best Local Similarity   80.0%;  Pred. No. 3.2e+03;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 GTGCTCCCTGTTTCATCAAT 20
           |||||| || |||| | |||
Db      29 GTGCTCACTTTTTCTTAAAT 10


RESULT 12
ABS18295/c
ID   ABS18295 standard; DNA; 94 BP.
XX
AC   ABS18295;
XX
DT   19-AUG-2002  (first entry)
XX
DE   Human genome-derived single exon probe ORF from lung SEQ ID No 18286.
XX
KW   Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;
KW   chronic obstructive pulmonary disease; interstitial lung disease;
KW   familial idiopathic pulmonary fibrosis; neurofibromatosis;
KW   tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
KW   Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
KW   pulmonary histiocytosis; lymphangioleiomyomtosis; Karagener syndrome;
KW   pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
KW   primary ciliary dyskinesis; pulmonary hypertension;
KW   hyaline membrane disease; open reading frame; ORF.
XX
OS   Homo sapiens.
XX
PN   WO200186003-A2.
XX
PD   15-NOV-2001.
XX
PF   30-JAN-2001; 2001WO-US00665.
XX
PR   04-FEB-2000; 2000US-180312P.
PR   26-MAY-2000; 2000US-207456P.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-234687P.
PR   27-SEP-2000; 2000US-236359P.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2002-114183/15.
XX
PT   Spatially-addressable set of single exon nucleic acid probes, used to
PT   measure gene expression in human lung samples -
XX
PS   Claim 4; SEQ ID No 18286; 634pp; English.
XX
CC   The invention relates to a spatially-addressable set of single exon
CC   nucleic acid probes for measuring gene expression in a sample derived
CC   from human lung comprising single exon nucleic acid probes having one of
CC   12614 nucleic acid sequences mentioned in the specification, or their
CC   complements or the 12387 open reading frames derived from the 12614
CC   probes. Also included are a microarray comprising the novel set of
CC   probes; the novel set of probes which hybridise at high stringency to a
CC   nucleic acid expressed in the human lung; measuring gene expression in a
CC   sample derived from human lung, comprising (a) contacting the array with
CC   a collection of detectably labeled nucleic acids derived from human lung
CC   mRNA, and (b) measuring the label detectably bound to each probe of
CC   the array; identifying exons in a eukaryotic genome, comprising
CC   (a) algorithmically predicting at least one exon from genomic sequences
CC   of the eukaryote; and (b) detecting specific hybridisation of detectably
CC   labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
CC   having a fragment identical to the predicted exon, the probe is included
CC   in the above mentioned microarray; assigning exons to a single gene,
CC   comprising (a) identifying exons from genomic sequence by the method
CC   above and (b) measuring the expression of each of the exons in several
CC   tissues and/or cell types using hybridisation to a single exon
CC   microarrays having a probe with the exon, where a common pattern of
CC   expression of the exons in the tissues and/or cell types indicates that
CC   the exons should be assigned to a single gene; a peptide comprising one
CC   of 12011 sequences, mentioned in the specification, or encoded by the
CC   probes/open reading frames (ORF). The probes are used for gene
CC   expression analysis, and for identifying exons in a gene, particularly
CC   using human lung derived mRNA and for the study of lung diseases
CC   such as asthma, lung cancer, chronic obstructive pulmonary disease
CC   (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
CC   fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,
CC   Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary
CC   haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomtosis,
CC   pulmonary alveolar proteinosis, Karagener syndrome, fibrocystic
CC   pulmonary dysplasia, primary ciliary dyskinesis, pulmonary hypertension
CC   and hyaline membrane disease. The present sequence is a single exon
CC   probe open reading frame of the invention.
CC   Note: The sequence data for this patent did not form part
CC   of the printed specification, but was obtained in electronic
CC   format directly from WIPO at
```

```
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 94 BP; 24 A; 11 C; 22 G; 37 T; 0 other;

  Query Match             68.0%;  Score 13.6;  DB 24;  Length 94;
  Best Local Similarity   80.0%;  Pred. No. 3.2e+03;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 GTGCTCCCTGTTTCATCAAT 20
           |||||| || |||| | |||
Db      29 GTGCTCACTTTTTCTTAAAT 10


RESULT 13
AAS50985
ID   AAS50985 standard; DNA; 99 BP.
XX
AC   AAS50985;
XX
DT   13-FEB-2002  (first entry)
XX
DE   Staphylococcus aureus cellular proliferation inhibitory sequence #2209.
XX
KW   Antisense; ss; prokaryotic cellular proliferation;
KW   antibiotic; antibacterial; drug design.
XX
OS   Staphylococcus aureus.
XX
PN   WO200170955-A2.
XX
PD   27-SEP-2001.
XX
PF   21-MAR-2001; 2001WO-US09180.
XX
PR   21-MAR-2000; 2000US-191078P.
PR   23-MAY-2000; 2000US-206848P.
PR   26-MAY-2000; 2000US-207727P.
PR   23-OCT-2000; 2000US-242578P.
PR   27-NOV-2000; 2000US-253625P.
PR   22-DEC-2000; 2000US-257931P.
PR   16-FEB-2001; 2001US-269308P.
XX
PA   (ELIT-) ELITRA PHARM INC.
XX
PI   Haselbeck R,  Ohlsen KL,  Zyskind JW,  Wall D,  Trawick JD,  Carr GJ;
PI   Yamamoto RT,  Xu HH;
XX
DR   WPI; 2001-611495/70.
XX
PT   New polynucleotides for the identification and development of
PT   antibiotics, comprise sequences of antisense nucleic acids -
XX
PS   Claim 1; Seq ID No 3562; 511pp; English.
XX
CC   The invention relates to antisense inhibitors of genes essential to
CC   prokaryotic cellular proliferation, their use in identifying the
CC   genes, their use in the discovery of novel antibiotics, the essential
CC   genes themselves and the encoded proteins. The prokaryotes used are
CC   Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC   pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC   invention is also useful for the identification of potential new targets
CC   for antibiotic development. The antisense nucleic acids can also be used
CC   to identify proteins used in proliferation, to express these proteins,
CC   and to obtain antibodies capable of binding to the expressed proteins.
CC   The proteins can be used to screen compounds in rational drug discovery
CC   programmes. The antisense nucleic acid sequence is also useful to screen
CC   for homologous nucleic acids which are required for cell proliferation in
CC   a wide variety of organisms. The present sequence is an antisense
CC   oligonucleotide of the invention.
CC   Note: The sequence data for this patent did not form part
CC   of the printed specification, but was obtained in electronic
CC   format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 99 BP; 25 A; 25 C; 23 G; 26 T; 0 other;

  Query Match             68.0%;  Score 13.6;  DB 23;  Length 99;
  Best Local Similarity   80.0%;  Pred. No. 3.2e+03;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 GTGCTCCCTGTTTCATCAAT 20
             ||||  || |||||||  |||
Db      14 GTGCAGCCGGTTTCATAAAT 33


RESULT 14
ABN46831/c
ID   ABN46831 standard; DNA; 60 BP.
XX
AC   ABN46831;
XX
DT   15-JUL-2002  (first entry)
XX
DE   Human spliced transcript detection oligonucleotide SEQ ID NO:19579.
XX
KW   Human; mouse; rat; splice transcript; detection; RNA transcript;
KW   splice variant; transcriptome; oligonucleotide library; ss.
XX
OS   Homo sapiens.
XX
PN   WO200210449-A2.
XX
PD   07-FEB-2002.
XX
PF   20-JUL-2001; 2001WO-IB01903.
XX
PR   28-JUL-2000; 2000US-221607P.
PR   02-MAY-2001; 2001US-287724P.
XX
PA   (COMP-) COMPUGEN INC.
XX
PI   Shoshan A,  Wasserman A,  Mintz E,  Mintz L,  Faigler S;
XX
DR   WPI; 2002-257383/30.
XX
PT   New oligonucleotide libraries comprising oligonucleotides which
PT   selectively hybridize to mRNAs transcribed from a transcription unit of
PT   a genome, useful for detecting tissue-, pathology-, and
PT   developmental-specific genes  -
XX
PS   Example 1; SEQ ID 19579; 47pp; English.
XX
CC   The present invention describes oligonucleotide libraries for detecting
CC   messenger RNAs that populate a (sub-)transcriptome, where the
CC   (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC   transcription units that populate a genome. The library comprises
CC   several oligonucleotides, each capable of hybridising selectively to a
CC   set of messenger RNAs transcribed from a given transcription unit of
CC   the genome, which encodes one or more messenger RNA splice variants.
CC   The oligonucleotide libraries are useful for detecting mRNAs from a
CC   biological sample, in expression profiling studies, in qualitatively or
CC   quantitatively characterising the corresponding transcriptome, and in
CC   detecting RNA transcripts and splice variants of human or animal
CC   transcriptomes. The libraries may also be used as specialised mini
CC   libraries to detect transcripts of a sub-transcriptome under a
CC   particular biological or pathological state, and so allowing the
CC   detection of tissue- and pathology-specific genes such as those genes
CC   only expressed in specific tissue under a specific pathological
CC   condition; to detect developmental specific genes; and to detect RNA
CC   transcripts and splice variants of a transcriptome of a patient suffering
CC   from a particular disorder. ABN27253 to ABN59589 represent
CC   oligonucleotide sequences from rats, humans and mice, which are used in
CC   the exemplification of the present invention.
CC   N.B. The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
```

```
XX
SQ   Sequence 60 BP; 19 A; 14 C; 10 G; 17 T; 0 other;

  Query Match             67.0%;  Score 13.4;  DB 24;  Length 60;
  Best Local Similarity   93.3%;  Pred. No. 3.8e+03;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       6 CCCTGTTTCATCAAT 20
           | ||||||||||||||
Db      46 CACTGTTTCATCAAT 32


RESULT 15
AAX85482
ID   AAX85482 standard; RNA; 93 BP.
XX
AC   AAX85482;
XX
DT   06-AUG-1999  (first entry)
XX
DE   Human artery perfusion selex RNA ligand SEQ ID NO:425.
XX
KW   Nucleic acid ligand; high-affinity; tissue target; blood vessel;
KW   atherosclerotic plaque; carotid artery; WHHL; diagnosis; therapy;
KW   watanabe heritable hyperlipidaemia; ss.
XX
OS   Homo sapiens.
XX
PN   WO9927138-A1.
XX
PD   03-JUN-1999.
XX
PF   19-NOV-1998;   98WO-US25006.
XX
PR   21-NOV-1997;   97US-0976413.
XX
PA   (NEXS-) NEXSTAR PHARM INC.
PA   (SCHD ) SCHERING AG.
XX
PI   Gold L,  Speck U,  Stephens A;
XX
DR   WPI; 1999-357856/30.
XX
PT   Identifying nucleic acid ligands to blood vessels
XX
PS   Claim 14; Page 78; 210pp; English.
XX
CC   The present invention describes a new method of identifying nucleic acid
CC   ligands to blood vessels. The method comprises contacting and
CC   partitioning nucleic acid sequences having increased affinity to the
CC   blood vessels and amplifying enriched sequences. The nucleic acid
CC   ligands are capable of binding specifically to tissues which are
CC   macromolecules in a heterogeneous environment, such as whole cells or
CC   substructures, aggregates of cells, collections of cells, or aggregates
CC   of macromolecules. The ligands can be used to identify and purify
CC   epitopes and macromolecules. The products can be used as diagnostic and
CC   therapeutic agents. AAX85058 to AAX85497 represent oligonucleotides used
CC   in the exemplification of the present invention.
XX
SQ   Sequence 93 BP; 29 A; 26 C; 20 G; 18 U; 0 other;

  Query Match             67.0%;  Score 13.4;  DB 20;  Length 93;
  Best Local Similarity   60.0%;  Pred. No. 4e+03;
  Matches    9;  Conservative    5;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       6 CCCTGTTTCATCAAT 20
           |||:|:::|| |||:
Db      38 CCCUGUUUCAACAAU 52


Search completed: November 23, 2002, 06:29:26
Job time : 102.6 secs
```

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 05:53:51 ; Search time 21.55 Seconds
(without alignments)
284.619 Million cell updates/sec

Title: US-09-296-264-18
Perfect score: 20
Sequence: 1 gtgctccctgtttcatcaat 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 687286

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	13.4	67.0	93	3	US-08-976-413A-425	Sequence 425, App
c	2	13.2	66.0	49	4	US-08-916-576B-26	Sequence 26, Appl
c	3	13	65.0	27	1	US-08-758-306-836	Sequence 836, App
c	4	13	65.0	27	4	US-08-584-040-4783	Sequence 4783, Ap
c	5	12.8	64.0	30	2	US-08-685-576-14	Sequence 14, Appl
c	6	12.8	64.0	31	4	US-08-679-645-429	Sequence 429, App
c	7	12.8	64.0	33	4	US-09-147-208-61	Sequence 61, Appl
c	8	12.6	63.0	30	1	US-08-479-817-9	Sequence 9, Appli
c	9	12.6	63.0	30	1	US-08-461-038-9	Sequence 9, Appli
c	10	12.6	63.0	30	1	US-08-461-645-9	Sequence 9, Appli
	11	12.6	63.0	37	2	US-08-874-678-14	Sequence 14, Appl
	12	12.6	63.0	37	3	US-08-643-839-14	Sequence 14, Appl
	13	12.6	63.0	37	4	US-09-348-886-14	Sequence 14, Appl
	14	12.6	63.0	38	3	US-08-894-784-48	Sequence 48, Appl
	15	12.6	63.0	41	2	US-08-702-572-8	Sequence 8, Appli
c	16	12.6	63.0	47	4	US-09-338-907-225	Sequence 225, App
c	17	12.6	63.0	47	4	US-09-338-907-302	Sequence 302, App
c	18	12.6	63.0	47	4	US-09-218-207-225	Sequence 225, App
c	19	12.6	63.0	47	4	US-09-218-207-302	Sequence 302, App
	20	12.6	63.0	63	2	US-08-874-678-41	Sequence 41, Appl
	21	12.6	63.0	63	3	US-08-643-839-41	Sequence 41, Appl
	22	12.6	63.0	63	4	US-09-348-886-41	Sequence 41, Appl
c	23	12.6	63.0	82	4	US-09-394-457C-16	Sequence 16, Appl
c	24	12.6	63.0	82	4	US-09-709-596A-16	Sequence 16, Appl
	25	12.6	63.0	87	4	US-09-394-457C-12	Sequence 12, Appl
	26	12.6	63.0	87	4	US-09-709-596A-12	Sequence 12, Appl
c	27	12.4	62.0	27	1	US-08-758-306-456	Sequence 456, App
c	28	12.4	62.0	27	3	US-08-985-162-1124	Sequence 1124, Ap
c	29	12.4	62.0	27	3	US-08-998-099-243	Sequence 243, App
c	30	12.4	62.0	27	4	US-08-584-040-427	Sequence 427, App
c	31	12.4	62.0	27	4	US-08-584-040-669	Sequence 669, App
c	32	12.4	62.0	27	4	US-08-584-040-1064	Sequence 1064, Ap
c	33	12.4	62.0	27	4	US-08-584-040-6629	Sequence 6629, Ap
c	34	12.4	62.0	27	4	US-08-584-040-6808	Sequence 6808, Ap
c	35	12.4	62.0	27	4	US-08-584-040-7079	Sequence 7079, Ap
	36	12.2	61.0	25	4	US-09-350-969-58	Sequence 58, Appl
c	37	12.2	61.0	27	1	US-08-758-306-620	Sequence 620, App
c	38	12.2	61.0	27	3	US-08-998-099-143	Sequence 143, App
c	39	12.2	61.0	54	4	US-08-679-645-620	Sequence 620, App
c	40	12	60.0	27	3	US-08-985-162-1427	Sequence 1427, Ap
c	41	12	60.0	27	3	US-08-985-162-1578	Sequence 1578, Ap
c	42	12	60.0	27	4	US-08-584-040-288	Sequence 288, App
c	43	12	60.0	27	4	US-08-584-040-6482	Sequence 6482, Ap
c	44	12	60.0	33	4	US-09-337-307A-8	Sequence 8, Appli
c	45	12	60.0	39	3	US-08-875-223-6	Sequence 6, Appli

ALIGNMENTS

RESULT 1
US-08-976-413A-425
; Sequence 425, Application US/08976413A
; Patent No. 6127119
; GENERAL INFORMATION:
; APPLICANT: STEPHENS, ANDREW
; APPLICANT: GOLD, LARRY
; APPLICANT: SPECK, ULRICH
; TITLE OF INVENTION: NUCLEIC ACID LIGANDS OF TISSUE TARGET
; NUMBER OF SEQUENCES: 440
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Swanson & Bratschun, L.L.C.
; STREET: 8400 E. Prentice Avenue, Suite 200
; CITY: Englewood
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3 1/2 diskette, 1.44 MG
; COMPUTER: IBM pc compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect 8.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/976,413A
; FILING DATE: 21-NOVEMBER-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/433,124
; FILING DATE: 03-MAY-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/714,131
; FILING DATE: 10-JUNE-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/536,428
; FILING DATE: 11-JUNE-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/964,624
; FILING DATE: 21-OCTOBER-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Barry J. Swanson
; REGISTRATION NUMBER: 33,215
; REFERENCE/DOCKET NUMBER: NEX31/CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 793-3333
; TELEFAX: (303) 793-3433
; INFORMATION FOR SEQ ID NO: 425:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 93
; TYPE: nucleic acid
; STRANDEDNESS: single

```
;      TOPOLOGY:  linear
;    FEATURE:
;      OTHER INFORMATION:  All C's are 2'-F cytosine
;    FEATURE:
;      OTHER INFORMATION:  All U's are 2'-F uracil
US-08-976-413A-425

  Query Match             67.0%;  Score 13.4;  DB 3;  Length 93;
  Best Local Similarity   60.0%;  Pred. No. 4.1e+02;
  Matches    9;  Conservative    5;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       6 CCCTGTTTCATCAAT 20
           |||:|:::|| |||:
Db      38 CCCUGUUUCAACAAU 52


RESULT 2
US-08-916-576B-26/c
; Sequence 26, Application US/08916576B
; Patent No. 6171816
;  GENERAL INFORMATION:
;    APPLICANT:  YU, GUO-LIANG
;    APPLICANT:  DILLON, PATRICK J.
;    APPLICANT:  EBNER, REINHARD
;    APPLICANT:  ENDRESS, GREGORY A.
;    TITLE OF INVENTION:  NOVEL HUMAN GROWTH FACTORS
;    NUMBER OF SEQUENCES:  45
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  STERNE, KESSLER, GOLDSTEIN & FOX, P.L.L.C.
;      STREET:  1100 NEW YORK AVENUE, SUITE 600
;      CITY:  WASHINGTON
;      STATE:  DC
;      COUNTRY:  US
;      ZIP:  20005-3934
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/916,576B
;      FILING DATE:
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/024,347
;      FILING DATE:  23-AUG-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  STEFFE, ERIC K.
;      REGISTRATION NUMBER:  36,688
;      REFERENCE/DOCKET NUMBER:  1488.0500001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 371-2600
;      TELEFAX:  (202) 371-2540
;  INFORMATION FOR SEQ ID NO:  26:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  49 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-916-576B-26

  Query Match             66.0%;  Score 13.2;  DB 4;  Length 49;
  Best Local Similarity   83.3%;  Pred. No. 4.7e+02;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 GTGCTCCCTGTTTCATCA 18
           |||| || | ||||||||
Db      33 GTGCCCCTTCTTTCATCA 16


RESULT 3
US-08-758-306-836/c
```

```
; Sequence 836, Application US/08758306
; Patent No. 5807743
;  GENERAL INFORMATION:
;    APPLICANT:  Stinchcomb, Dan T.
;    APPLICANT:  McSwiggen, James A.
;    TITLE OF INVENTION:  METHOD AND REAGENT FOR THE
;    TITLE OF INVENTION:  TREATMENT OF DISEASES
;    TITLE OF INVENTION:  ASSOCIATED WITH
;    TITLE OF INVENTION:  INTERLEUKIN-2 RECEPTOR
;    TITLE OF INVENTION:  GAMMA-CHAIN EXPRESSION
;    NUMBER OF SEQUENCES:  1379
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Lyon & Lyon
;      STREET:  633 West Fifth Street
;      STREET:  Suite 4700
;      CITY:  Los Angeles
;      STATE:  California
;      COUNTRY:  U.S.A.
;      ZIP:  90071-2066
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" Diskette, 1.44 Mb
;      MEDIUM TYPE:  storage
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  IBM P.C. DOS 5.0
;      SOFTWARE:  FastSeq Version 1.5
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/758,306
;      FILING DATE:  December 3, 1996
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Warburg, Richard J.
;      REGISTRATION NUMBER:  32,327
;      REFERENCE/DOCKET NUMBER:  212/132
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (213) 489-1600
;      TELEFAX:  (213) 955-0440
;      TELEX:  67-3510
;  INFORMATION FOR SEQ ID NO:  836:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  27 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    FEATURE:
;      OTHER INFORMATION:  The letter "N" stands for the stem II
;      OTHER INFORMATION:  region of a HH ribozyme.
US-08-758-306-836

  Query Match             65.0%;  Score 13;  DB 1;  Length 27;
  Best Local Similarity   92.9%;  Pred. No. 5.4e+02;
  Matches   13;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       5 TCCCTGTTTCATCA 18
           |||||||||| |||
Db      26 TCCCTGTTTCNTCA 13


RESULT 4
US-08-584-040-4783/c
; Sequence 4783, Application US/08584040
; Patent No. 6346398
;  GENERAL INFORMATION:
;    APPLICANT:  Pavco, Pamela
;    APPLICANT:  McSwiggen, James
;    APPLICANT:  Stinchcomb, Dan T.
;    APPLICANT:  Escobedo, Jaime
;    TITLE OF INVENTION:  METHOD AND REAGENT FOR THE
;    TITLE OF INVENTION:  TREATMENT OF DISEASES OR
;    TITLE OF INVENTION:  CONDITIONS RELATED TO LEVELS
;    TITLE OF INVENTION:  OF VASCULAR ENDOTHELIAL
```

```
;    TITLE OF INVENTION:  GROWTH FACTOR
;    NUMBER OF SEQUENCES:  8502
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Lyon & Lyon
;      STREET:  633 West Fifth Street
;      STREET:  Suite 4700
;      CITY:  Los Angeles
;      STATE:  California
;      COUNTRY:  U.S.A.
;      ZIP:  90071-2066
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" Diskette, 1.44 Mb
;      MEDIUM TYPE:  storage
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  IBM P.C. DOS 5.0
;      SOFTWARE:  Word Perfect 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/584,040
;      FILING DATE:  January 11, 1996
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/005,974
;      FILING DATE:  October 26, 1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Warburg, Richard J.
;      REGISTRATION NUMBER:  32,327
;      REFERENCE/DOCKET NUMBER:  218/064
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (213) 489-1600
;      TELEFAX:  (213) 955-0440
;      TELEX:  67-3510
;  INFORMATION FOR SEQ ID NO:  4783:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  27 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    FEATURE:
;      OTHER INFORMATION:  The letter "N" represents the stem II region
;      OTHER INFORMATION:  of an HH ribozyme.
US-08-584-040-4783

  Query Match             65.0%;  Score 13;  DB 4;  Length 27;
  Best Local Similarity   92.9%;  Pred. No. 5.4e+02;
  Matches   13;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       5 TCCCTGTTTCATCA 18
           |||||||||| |||
Db      26 TCCCTGTTTCNTCA 13


RESULT 5
US-08-685-576-14/c
; Sequence 14, Application US/08685576
; Patent No. 5906819
;  GENERAL INFORMATION:
;    APPLICANT:  Kaibuchi, Kozo
;    APPLICANT:  Iwamatsu, Akihiro
;    APPLICANT:  Nakano, Takeshi
;    APPLICANT:  Ito, Masaaki
;    APPLICANT:  Takahashi, No. 5906819uaki
;    TITLE OF INVENTION:  RHO TARGET PROTEIN RHO-KINASE
;    NUMBER OF SEQUENCES:  16
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  3000 K Street, N.W., Suite 500
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20007-5109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/685,576
;      FILING DATE:  24-JUL-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 7-325129
;      FILING DATE:  20-NOV-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 8-17150
;      FILING DATE:  05-JAN-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 8-131206
;      FILING DATE:  26-APR-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Bent, Stephen A.
;      REGISTRATION NUMBER:  29,768
;      REFERENCE/DOCKET NUMBER:  16887/843
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)672-5300
;      TELEFAX:  (202)672-5399
;      TELEX:  904136
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  30 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  other nucleic acid
;      DESCRIPTION:  /desc = "primer"
US-08-685-576-14

  Query Match             64.0%;  Score 12.8;  DB 2;  Length 30;
  Best Local Similarity   87.5%;  Pred. No. 7e+02;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       5 TCCCTGTTTCATCAAT 20
           | |||||||||||  ||
Db      27 TGCCTGTTTCATCCAT 12


RESULT 6
US-08-679-645-429/c
; Sequence 429, Application US/08679645
; Patent No. 6350934
;  GENERAL INFORMATION:
;    APPLICANT:  Zwick, Michael G.
;    APPLICANT:  Edington, Brent E.
;    APPLICANT:  McSwiggen, James A.
;    APPLICANT:  Merlo, Patricia Ann Owens
;    APPLICANT:  Guo, Lining
;    APPLICANT:  Skokut, Thomas A.
;    APPLICANT:  Young, Scott A.
;    APPLICANT:  Folkerts, Otto
;    APPLICANT:  Merlo, Donald J.
;    TITLE OF INVENTION:  COMPOSITION AND METHODS FOR
;    TITLE OF INVENTION:  MODULATION OF GENE EXPRESSION
;    TITLE OF INVENTION:  IN PLANTS
;    NUMBER OF SEQUENCES:  1263
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Lyon & Lyon
;      STREET:  633 West Fifth Street
;      STREET:  Suite 4700
;      CITY:  Los Angeles
;      STATE:  California
;      COUNTRY:  U.S.A.
;      ZIP:  90071-2066
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5" Diskette, 1.44 Mb
;      MEDIUM TYPE:  storage
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  IBM P.C. DOS 5.0
```

```
;      SOFTWARE:  Word Perfect 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/679,645
;      FILING DATE:  July 12, 1996
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/001,135
;      FILING DATE:  July 13, 1995
;      APPLICATION NUMBER:  08/300,726
;      FILING DATE:  September 2, 1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Warburg, Richard J.
;      REGISTRATION NUMBER:  32,327
;      REFERENCE/DOCKET NUMBER:  219/247
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (213) 489-1600
;      TELEFAX:  (213) 955-0440
;      TELEX:  67-3510
;  INFORMATION FOR SEQ ID NO:  429:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  31 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    FEATURE:
;      OTHER INFORMATION:  The letter "N" stands for any base.
US-08-679-645-429

  Query Match             64.0%;  Score 12.8;  DB 4;  Length 31;
  Best Local Similarity   82.4%;  Pred. No. 7e+02;
  Matches   14;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 TGCTCCCTGTTTCATCA 18
           ||||| | ||||| |||
Db      31 TGCTCGCCGTTTCNTCA 15


RESULT 7
US-09-147-208-61/c
; Sequence 61, Application US/09147208
; Patent No. 6333303
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  Antiviral Ricin-Like Proteins
;    NUMBER OF SEQUENCES:  71
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  BERESKIN & PARR
;      STREET:  40 King Street West
;      CITY:  Toronto
;      STATE:  Ontario
;      COUNTRY:  Canada
;      ZIP:  M5H 3Y2
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/147,208
;      FILING DATE:  02-MAR-1999
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Rudolph, John R.
;      REGISTRATION NUMBER:  38,003
;      REFERENCE/DOCKET NUMBER:  7841-76
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (416) 364-7311
;      TELEFAX:  (416) 361-1398
;  INFORMATION FOR SEQ ID NO:  61:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  33 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-09-147-208-61

  Query Match             64.0%;  Score 12.8;  DB 4;  Length 33;
  Best Local Similarity   87.5%;  Pred. No. 7e+02;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       2 TGCTCCCTGTTTCATC 17
           | ||||| ||||||||
Db      19 TCCTCCCGGTTTCATC 4


RESULT 8
US-08-479-817-9/c
; Sequence 9, Application US/08479817
; Patent No. 5597910
;  GENERAL INFORMATION:
;    APPLICANT:  Gudibande, Satyanarayana R.
;    APPLICANT:  Kenten, John H.
;    TITLE OF INVENTION:  IMPROVED ELECTROCHEMILUMINESCENT LABEL
;    TITLE OF INVENTION:  FOR DNA PROBE ASSAYS
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Curtis, Morris & Safford
;      ADDRESSEE:  c/o Barry Evans
;      STREET:  530 Fifth Avenue
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/479,817
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/307,026
;      FILING DATE:
;      APPLICATION NUMBER:  US 07/805,537
;      FILING DATE:  11-DEC-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Evans, Barry
;      REGISTRATION NUMBER:  22,802
;      REFERENCE/DOCKET NUMBER:  370068-3440
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 840-3333
;      TELEFAX:  (212) 840-0712
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  30 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-479-817-9

  Query Match             63.0%;  Score 12.6;  DB 1;  Length 30;
  Best Local Similarity   78.9%;  Pred. No. 8.8e+02;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 GTGCTCCCTGTTTCATCAA 19
           ||||| ||||| |||  ||
Db      19 GTGCTACCTGTGTCACAAA 1


RESULT 9
US-08-461-038-9/c
; Sequence 9, Application US/08461038
; Patent No. 5610017
```

```
;  GENERAL INFORMATION:
;    APPLICANT:  Gudibande, Satyanarayana R.
;    APPLICANT:  Kenten, John H.
;    TITLE OF INVENTION:  IMPROVED ELECTROCHEMILUMINESCENT LABEL
;    TITLE OF INVENTION:  FOR DNA PROBE ASSAYS
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Curtis, Morris & Safford
;      ADDRESSEE:  c/o Barry Evans
;      STREET:  530 Fifth Avenue
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/461,038
;      FILING DATE:  05-JUN-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Evans, Barry
;      REGISTRATION NUMBER:  22,802
;      REFERENCE/DOCKET NUMBER:  370068-3451
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 840-3333
;      TELEFAX:  (212) 840-0712
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  30 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-461-038-9

  Query Match             63.0%;  Score 12.6;  DB 1;  Length 30;
  Best Local Similarity   78.9%;  Pred. No. 8.8e+02;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 GTGCTCCCTGTTTCATCAA 19
          ||||| ||||| |||  ||
Db     19 GTGCTACCTGTGTCACAAA 1


RESULT 10
US-08-461-645-9/c
; Sequence 9, Application US/08461645
; Patent No. 5686244
;  GENERAL INFORMATION:
;    APPLICANT:  Gudibande, Satyanarayana R.
;    APPLICANT:  Kenten, John H.
;    TITLE OF INVENTION:  IMPROVED ELECTROCHEMILUMINESCENT LABEL
;    TITLE OF INVENTION:  FOR DNA PROBE ASSAYS
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Curtis, Morris & Safford
;      ADDRESSEE:  c/o Barry Evans
;      STREET:  530 Fifth Avenue
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10036
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/461,645
;      FILING DATE:  05-JUN-1995
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Evans, Barry
;      REGISTRATION NUMBER:  22,802
;      REFERENCE/DOCKET NUMBER:  370068-3450
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 840-3333
;      TELEFAX:  (212) 840-0712
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  30 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-461-645-9

  Query Match             63.0%;  Score 12.6;  DB 1;  Length 30;
  Best Local Similarity   78.9%;  Pred. No. 8.8e+02;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 GTGCTCCCTGTTTCATCAA 19
          ||||| ||||| |||  ||
Db     19 GTGCTACCTGTGTCACAAA 1


RESULT 11
US-08-874-678-14
; Sequence 14, Application US/08874678
; Patent No. 5952199
;  GENERAL INFORMATION:
;    APPLICANT:  Davis-Smyth, Terri L.
;    APPLICANT:  Chen, Helen H.
;    APPLICANT:  Presta, Leonard
;    APPLICANT:  Ferrara, Napoleone
;    TITLE OF INVENTION:  NOVEL INHIBITORS OF VASCULAR ENDOTHELIAL GROWTH FACTOR
;    TITLE OF INVENTION:  ACTIVITY, THEIR USES AND PROCESSES FOR THEIR PRODUCTION
;    NUMBER OF SEQUENCES:  48
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Flehr, Hohbach, Test, Albritton & Herbert
;      STREET:  Four Embarcadero Center, Suite 3400
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  United States
;      ZIP:  94111-4187
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/874,678
;      FILING DATE:  HEREWITH
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/643,839
;      FILING DATE:  07-MAY-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Dreger, Walter H.
;      REGISTRATION NUMBER:  24,190
;      REFERENCE/DOCKET NUMBER:  A-63291-1/WHD
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 781-1989
;      TELEFAX:  (415) 398-3249
;      TELEX:  910 277299
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  37 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  unknown
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  CDS
```

```
;      LOCATION:  1..36
US-08-874-678-14

  Query Match             63.0%;  Score 12.6;  DB 2;  Length 37;
  Best Local Similarity   78.9%;  Pred. No. 9e+02;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      2 TGCTCCCTGTTTCATCAAT 20
          || | || |||||||||  |
Db     12 TGATACCGGTTTCATCACT 30


RESULT 12
US-08-643-839-14
; Sequence 14, Application US/08643839
; Patent No. 6100071
;  GENERAL INFORMATION:
;    APPLICANT:  Davis-Smyth, Terri L.
;    APPLICANT:  Chen, Helen H.
;    APPLICANT:  Presta, Leonard
;    APPLICANT:  Ferrara, Napoleone
;    TITLE OF INVENTION:  NOVEL INHIBITORS OF VASCULAR ENDOTHELIAL
;    TITLE OF INVENTION:  GROWTH FACTOR ACTIVITY, THEIR USES AND PROCESSES FOR THEIR
;    TITLE OF INVENTION:  PRODUCTION
;    NUMBER OF SEQUENCES:  48
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Flehr, Hohbach, Test, Albritton & Herbert
;      STREET:  Four Embarcadero Center, Suite 3400
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  United States
;      ZIP:  94111-4187
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/643,839
;      FILING DATE:  07-MAY-1996
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Dreger, Walter H.
;      REGISTRATION NUMBER:  24,190
;      REFERENCE/DOCKET NUMBER:  A-63291/WHD
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 781-1989
;      TELEFAX:  (415) 398-3249
;      TELEX:  910 277299
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  37 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  unknown
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  1..36
US-08-643-839-14

  Query Match             63.0%;  Score 12.6;  DB 3;  Length 37;
  Best Local Similarity   78.9%;  Pred. No. 9e+02;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      2 TGCTCCCTGTTTCATCAAT 20
          || | || |||||||||  |
Db     12 TGATACCGGTTTCATCACT 30


RESULT 13
US-09-348-886-14
; Sequence 14, Application US/09348886
; Patent No. 6383486
;  GENERAL INFORMATION:
;    APPLICANT:  Davis-Smyth, Terri L.
;    APPLICANT:  Chen, Helen H.
;    APPLICANT:  Presta, Leonard
;    APPLICANT:  Ferrara, Napoleone
;    TITLE OF INVENTION:  NOVEL INHIBITORS OF VASCULAR ENDOTHELIAL
;    TITLE OF INVENTION:  GROWTH FACTOR ACTIVITY, THEIR USES AND PROCESSES FOR THEIR
;    TITLE OF INVENTION:  PRODUCTION
;    NUMBER OF SEQUENCES:  48
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Flehr Hohbach Test Albritton & Herbert LLP
;      STREET:  Four Embarcadero Center, Suite 3400
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  United States
;      ZIP:  94111-4187
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/348,886
;      FILING DATE:  01-JUL-1999
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/643,839
;      FILING DATE:  07-MAY-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Dolly A. Vance
;      REGISTRATION NUMBER:  39,054
;      REFERENCE/DOCKET NUMBER:  A-63291-2/RMS/DAV
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 781-1989
;      TELEFAX:  (415) 398-3249
;      TELEX:  910 277299
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  37 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  unknown
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  DNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  1..36
US-09-348-886-14

  Query Match             63.0%;  Score 12.6;  DB 4;  Length 37;
  Best Local Similarity   78.9%;  Pred. No. 9e+02;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      2 TGCTCCCTGTTTCATCAAT 20
          || | || |||||||||  |
Db     12 TGATACCGGTTTCATCACT 30


RESULT 14
US-08-894-784-48
; Sequence 48, Application US/08894784
; Patent No. 6005095
;  GENERAL INFORMATION:
;    APPLICANT:  Capaccioli, Sergio
;    APPLICANT:  Morelli, Susanna
;    APPLICANT:  Nicolin, Angelo
;    TITLE OF INVENTION:  ANTISENSE TRANSCRIPT ASSOCIATED TO TUMOR
;    TITLE OF INVENTION:  CELLS HAVING A T(14:18) TRANSLOCATION AND
;    TITLE OF INVENTION:  OLIGODEOXYNUCLEOTIDES USEFUL IN THE DIAGNOSIS AND
;    TITLE OF INVENTION:  TREATMENT OF SAID TUMOR CELLS
;    NUMBER OF SEQUENCES:  49
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  FINNEGAN, HENDERSON, FARABOW, GARRETT &
```

```
;      ADDRESSEE:  DUNNER, LLP
;      STREET:  1300 I Street, NW
;      CITY:  Washington
;      STATE:  DC
;      COUNTRY:  USA
;      ZIP:  20005-3315
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/894,784
;      FILING DATE:  15-DEC-1997
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/EP96/00852
;      FILING DATE:  02-MAR-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  IL MI95 A 000420
;      FILING DATE:  03-MAR-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Paul, John C.
;      REGISTRATION NUMBER:  30,413
;      REFERENCE/DOCKET NUMBER:  05999.0005-00000
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-408-4000
;      TELEFAX:  202-408-4400
;  INFORMATION FOR SEQ ID NO:  48:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  38 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA
US-08-894-784-48

  Query Match             63.0%;  Score 12.6;  DB 3;  Length 38;
  Best Local Similarity   78.9%;  Pred. No. 9.1e+02;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 GTGCTCCCTGTTTCATCAA 19
           |||||||||| |||  | |
Db       8 GTGCTCCCTGGTTCCCCGA 26


RESULT 15
US-08-702-572-8
; Sequence 8, Application US/08702572
; Patent No. 5965386
;  GENERAL INFORMATION:
;    APPLICANT:  Kerry-Williams, Sean M
;    APPLICANT:  Gilbert, Sarah C
;    TITLE OF INVENTION:  Yeast Strains and Modified Albumins
;    NUMBER OF SEQUENCES:  16
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Centeon L.L.C.
;      STREET:  1020 First Avenue
;      CITY:  King of Prussia
;      STATE:  Pennsylvania
;      COUNTRY:  USA
;      ZIP:  19406-1310
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  MicroSoft Word 6.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/702,572
;      FILING DATE:  11-NOV-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  WO 95/23857
;      FILING DATE:  1-MAR-1995
;      APPLICATION NUMBER:  GB 9404270.2
;      FILING DATE:  5-MAR-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Naomi Biswas
;      REGISTRATION NUMBER:  38,384
;      REFERENCE/DOCKET NUMBER:  CE0114 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  610/878/4294
;      TELEFAX:  610/878/4221
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  41 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  YES
US-08-702-572-8

  Query Match             63.0%;  Score 12.6;  DB 2;  Length 41;
  Best Local Similarity   78.9%;  Pred. No. 9.2e+02;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 TGCTCCCTGTTTCATCAAT 20
           || | ||||| |||||| |
Db       8 TGGTACCTGTGTCATCACT 26


Search completed: November 23, 2002, 06:36:23
Job time : 23.55 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 05:54:41 ; Search time 17.25 Seconds
(without alignments)
439.108 Million cell updates/sec

Title: US-09-296-264-18
Perfect score: 20
Sequence: 1 gtgctccctgtttcatcaat 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 335578 seqs, 189365133 residues

Total number of hits satisfying chosen parameters: 195614

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	14.4	72.0	52	10	US-09-878-574-15196	Sequence 15196, A
c	2	13.6	68.0	94	10	US-09-864-761-22117	Sequence 22117, A
	3	13.6	68.0	99	10	US-09-815-242-3562	Sequence 3562, Ap
c	4	13.2	66.0	24	10	US-09-996-630-8	Sequence 8, Appli
	5	12.8	64.0	29	10	US-09-729-674-217	Sequence 217, App
c	6	12.6	63.0	24	9	US-09-903-327A-22	Sequence 22, Appl
	7	12.6	63.0	34	10	US-09-757-207-25	Sequence 25, Appl
	8	12.6	63.0	36	10	US-09-757-207-23	Sequence 23, Appl
c	9	12.6	63.0	47	9	US-09-853-526-225	Sequence 225, App
c	10	12.6	63.0	47	9	US-09-853-526-302	Sequence 302, App
c	11	12.6	63.0	47	10	US-09-901-484A-225	Sequence 225, App
c	12	12.6	63.0	47	10	US-09-901-484A-302	Sequence 302, App
	13	12.4	62.0	18	10	US-09-969-373-3369	Sequence 3369, Ap
c	14	12.4	62.0	18	10	US-09-969-373-3370	Sequence 3370, Ap
c	15	12.4	62.0	73	10	US-09-969-373-919	Sequence 919, App
c	16	12.2	61.0	20	10	US-09-131-827A-8	Sequence 8, Appli
c	17	12.2	61.0	26	10	US-09-837-235-35	Sequence 35, Appl
	18	12.2	61.0	26	10	US-09-837-235-36	Sequence 36, Appl
	19	12	60.0	65	10	US-09-878-574-10483	Sequence 10483, A
	20	12	60.0	83	10	US-09-864-761-23074	Sequence 23074, A
c	21	12	60.0	85	10	US-09-294-093B-4358	Sequence 4358, Ap
	22	12	60.0	98	10	US-09-864-761-26028	Sequence 26028, A
	23	11.8	59.0	20	10	US-09-755-004-6	Sequence 6, Appli
c	24	11.8	59.0	25	9	US-09-230-926A-8	Sequence 8, Appli
c	25	11.8	59.0	75	10	US-09-864-761-30517	Sequence 30517, A
	26	11.8	59.0	78	10	US-09-969-373-872	Sequence 872, App
	27	11.6	58.0	19	9	US-09-853-526-542	Sequence 542, App
	28	11.6	58.0	19	10	US-09-901-484A-542	Sequence 542, App
	29	11.6	58.0	24	10	US-09-898-779-12	Sequence 12, Appl
	30	11.6	58.0	24	10	US-09-898-779-13	Sequence 13, Appl
c	31	11.6	58.0	31	10	US-09-798-042-79	Sequence 79, Appl
c	32	11.6	58.0	32	10	US-09-798-042-83	Sequence 83, Appl
c	33	11.6	58.0	42	10	US-09-101-807-5	Sequence 5, Appli
c	34	11.6	58.0	68	10	US-09-878-574-12763	Sequence 12763, A
	35	11.6	58.0	87	10	US-09-893-737-135	Sequence 135, App
c	36	11.6	58.0	89	10	US-09-864-761-27490	Sequence 27490, A
c	37	11.6	58.0	97	10	US-09-864-761-25041	Sequence 25041, A
	38	11.4	57.0	22	12	US-10-028-415-28	Sequence 28, Appl
c	39	11.4	57.0	23	10	US-09-216-393-229	Sequence 229, App
	40	11.4	57.0	23	10	US-09-811-259-10	Sequence 10, Appl
c	41	11.4	57.0	59	12	US-10-013-737-8	Sequence 8, Appli
c	42	11.4	57.0	78	10	US-09-864-761-27522	Sequence 27522, A
c	43	11.4	57.0	81	10	US-09-864-761-32980	Sequence 32980, A
c	44	11.4	57.0	100	10	US-09-864-761-20823	Sequence 20823, A
	45	11.2	56.0	20	10	US-09-756-910-21	Sequence 21, Appl

ALIGNMENTS

RESULT 1
US-09-878-574-15196/c
; Sequence 15196, Application US/09878574
; Patent No. US20020110548A1
; GENERAL INFORMATION:
; APPLICANT: Byrum, Joseph R.
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Thompson, Michael D.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(15401)B
; CURRENT APPLICATION NUMBER: US/09/878,574
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 09/333,535
; PRIOR FILING DATE: 1999-06-14
; NUMBER OF SEQ ID NOS: 15775
; SEQ ID NO 15196
; LENGTH: 52
; TYPE: DNA
; ORGANISM: Glycine max
; OTHER INFORMATION: Clone ID: 701069776H1
US-09-878-574-15196

Query Match 72.0%; Score 14.4; DB 10; Length 52;
Best Local Similarity 93.8%; Pred. No. 2.5e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy       4 CTCCCTGTTTCATCAA 19
           |||| ||||||||||
Db      38 CTCCATGTTTCATCAA 23
```

RESULT 2
US-09-864-761-22117/c
; Sequence 22117, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO

```
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 22117
;   LENGTH: 94
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AC006010.2
;   OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.5
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.4
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.4
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.6
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.3
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.4
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.5
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.5
;   OTHER INFORMATION: SWISSPROT HIT: P46589, EVALUE 1.10e+00
;   OTHER INFORMATION: EST_HUMAN HIT: AA525207.1, EVALUE 1.10e+00
;   OTHER INFORMATION: NT HIT: X56600.1, EVALUE 2.00e-01
US-09-864-761-22117

  Query Match             68.0%;  Score 13.6;  DB 10;  Length 94;
  Best Local Similarity   80.0%;  Pred. No. 6.8e+02;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 GTGCTCCCTGTTTCATCAAT 20
           |||||| || |||| | |||
Db      29 GTGCTCACTTTTTCTTAAAT 10


RESULT 3
US-09-815-242-3562
; Sequence 3562, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
;  APPLICANT: Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari L.
;  APPLICANT:  Zyskind, Judith W.
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John D.
;  APPLICANT:  Carr, Grant J.
;  APPLICANT:  Yamamoto, Robert T.
;  APPLICANT:  Xu, H. Howard
;  TITLE OF INVENTION: Identification of Essential Genes in
;  TITLE OF INVENTION:  Prokaryotes
;  FILE REFERENCE: ELITRA.011A
;  CURRENT APPLICATION NUMBER: US/09/815,242
;  CURRENT FILING DATE:  2001-03-21
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  NUMBER OF SEQ ID NOS: 14110
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3562
;   LENGTH: 99
;   TYPE: DNA
;   ORGANISM: Staphylococcus aureus
US-09-815-242-3562

  Query Match             68.0%;  Score 13.6;  DB 10;  Length 99;
  Best Local Similarity   80.0%;  Pred. No. 6.9e+02;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 GTGCTCCCTGTTTCATCAAT 20
           ||||  || |||||| |||
Db      14 GTGCAGCCGGTTTCATAAAT 33


RESULT 4
US-09-996-630-8/c
; Sequence 8, Application US/09996630
; Patent No. US20020115090A1
; GENERAL INFORMATION:
;  APPLICANT: Gillis, Kimberly
;  APPLICANT:  Zhang, Yixian
;  TITLE OF INVENTION: Expression Analysis of KIAA Nucleic Acids And Polypeptides
;  TITLE OF INVENTION:  Useful In The Diagnosis And Treatment of Prostate Cancer
;  FILE REFERENCE: 102729-10
;  CURRENT APPLICATION NUMBER: US/09/996,630
;  CURRENT FILING DATE:  2002-06-18
;  PRIOR APPLICATION NUMBER: 60/253,460
;  PRIOR FILING DATE: 2000-11-28
;  NUMBER OF SEQ ID NOS: 9
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 8
;   LENGTH: 24
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-996-630-8

  Query Match             66.0%;  Score 13.2;  DB 10;  Length 24;
  Best Local Similarity   83.3%;  Pred. No. 8.5e+02;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 GTGCTCCCTGTTTCATCA 18
```

```
          ||| || ||||||| |||
Db     18 GTGGTCACTGTTTCTTCA 1


RESULT 5
US-09-729-674-217
; Sequence 217, Application US/09729674
; Patent No. US20010039335A1
; GENERAL INFORMATION:
;  APPLICANT: Jacobs, Kenneth
;  APPLICANT:  McCoy, John M.
;  APPLICANT:  LaVallie, Edward R.
;  APPLICANT:  Collins-Racie, Lisa A.
;  APPLICANT:  Evans, Cheryl
;  APPLICANT:  Merberg, David
;  APPLICANT:  Treacy, Maurice
;  APPLICANT:  Agostino, Michael J.
;  APPLICANT:  Steininger II, Robert J.
;  APPLICANT:  Spaulding, Vikki
;  APPLICANT:  Wong, Gordon G.
;  APPLICANT:  Clark, Hilary
;  APPLICANT:  Fechtel, Kim
;  APPLICANT:  Genetics Institute, Inc.
;  TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
;  FILE REFERENCE: 6055-64X
;  CURRENT APPLICATION NUMBER: US/09/729,674
;  CURRENT FILING DATE:  2000-12-04
;  PRIOR APPLICATION NUMBER: 09/539,330
;  PRIOR FILING DATE: 2000-03-30
;  NUMBER OF SEQ ID NOS: 283
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 217
;   LENGTH: 29
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: oligonucleotide
;   NAME/KEY: misc_feature
;   LOCATION: (2)
;   OTHER INFORMATION: biotinylated phosphoaramidite residue
US-09-729-674-217

  Query Match             64.0%;  Score 12.8;  DB 10;  Length 29;
  Best Local Similarity   87.5%;  Pred. No. 1.4e+03;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       3 GCTCCCTGTTTCATCA 18
           ||| || |||||||||
Db      12 GCTGCCCGTTTCATCA 27


RESULT 6
US-09-903-327A-22/c
; Sequence 22, Application US/09903327A
; Patent No. US20020164333A1
; GENERAL INFORMATION:
;  APPLICANT: Nemerow, Glen R.
;  APPLICANT:  Li, Erguang
;  TITLE OF INVENTION: BIFUNCTIONAL MOLECULES AND VECTORS COMPLEXED THEREWITH FOR TARGET
;  TITLE OF INVENTION:  GENE
;  TITLE OF INVENTION:  DELIVERY
;  FILE REFERENCE: 22908-1228
;  CURRENT APPLICATION NUMBER: US/09/903,327A
;  CURRENT FILING DATE:  2001-07-10
;  PRIOR APPLICATION NUMBER: 09/613,017
;  PRIOR FILING DATE: 2000-07-10
;  NUMBER OF SEQ ID NOS: 33
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 22
;   LENGTH: 24
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: PCR antisense primer for subcloning DAV-1 heavy chain for
;   OTHER INFORMATION: whole antibody construct.
US-09-903-327A-22

  Query Match             63.0%;  Score 12.6;  DB 9;  Length 24;
  Best Local Similarity   78.9%;  Pred. No. 1.7e+03;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 TGCTCCCTGTTTCATCAAT 20
           |||||  |||| ||| |||
Db      21 TGCTCTGTGTTACATGAAT 3


RESULT 7
US-09-757-207-25
; Sequence 25, Application US/09757207
; Patent No. US20020150880A1
; GENERAL INFORMATION:
;  APPLICANT:  Hellyer, Tobin J.
;  APPLICANT:  You, Qimin
;  APPLICANT:  Harris, James M.
;  TITLE OF INVENTION:  Sequences and Methods for Detection of HIV-1
;  FILE REFERENCE:  Seq/Mtds for Detection of HIV-1
;  CURRENT APPLICATION NUMBER: US/09/757,207
;  CURRENT FILING DATE:  2001-01-09
;  NUMBER OF SEQ ID NOS:  37
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 25
;   LENGTH: 34
;   TYPE: DNA
;   ORGANISM: Human immunodeficiency virus type 1
US-09-757-207-25

  Query Match             63.0%;  Score 12.6;  DB 10;  Length 34;
  Best Local Similarity   78.9%;  Pred. No. 1.8e+03;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 GTGCTCCCTGTTTCATCAA 19
           ||||||||| ||  || ||
Db      11 GTGCTCCCTTTTCTATTAA 29


RESULT 8
US-09-757-207-23
; Sequence 23, Application US/09757207
; Patent No. US20020150880A1
; GENERAL INFORMATION:
;  APPLICANT:  Hellyer, Tobin J.
;  APPLICANT:  You, Qimin
;  APPLICANT:  Harris, James M.
;  TITLE OF INVENTION:  Sequences and Methods for Detection of HIV-1
;  FILE REFERENCE:  Seq/Mtds for Detection of HIV-1
;  CURRENT APPLICATION NUMBER: US/09/757,207
;  CURRENT FILING DATE:  2001-01-09
;  NUMBER OF SEQ ID NOS:  37
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 23
;   LENGTH: 36
;   TYPE: DNA
;   ORGANISM: Human immunodeficiency virus type 1
US-09-757-207-23

  Query Match             63.0%;  Score 12.6;  DB 10;  Length 36;
  Best Local Similarity   78.9%;  Pred. No. 1.8e+03;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 GTGCTCCCTGTTTCATCAA 19
           ||||||||| ||  || ||
Db      11 GTGCTCCCTTTTCTATTAA 29


RESULT 9
US-09-853-526-225/c
```

```
; Sequence 225, Application US/09853526
; Patent No. US20020165345A1
; GENERAL INFORMATION:
;  APPLICANT: Cohen, Daniel
;  APPLICANT:  Blumenfeld, Marta
;  APPLICANT:  Ilya, Chumakov
;  APPLICANT:  Bougueleret, Lydie
;  TITLE OF INVENTION: PROSTATE CANCER GENE
;  FILE REFERENCE: GENSET.18CP1CP
;  CURRENT APPLICATION NUMBER: US/09/853,526
;  CURRENT FILING DATE:  2001-05-11
;  PRIOR APPLICATION NUMBER: 09/338,907
;  PRIOR FILING DATE: 1999-06-23
;  PRIOR APPLICATION NUMBER: 08/996,306
;  PRIOR FILING DATE: 1997-12-22
;  PRIOR APPLICATION NUMBER: 60/099,658
;  PRIOR FILING DATE: 1998-09-09
;  PRIOR APPLICATION NUMBER: 09/218,207
;  PRIOR FILING DATE: 1998-12-22
;  NUMBER OF SEQ ID NOS: 578
;  SOFTWARE: Patent.pm
; SEQ ID NO 225
;   LENGTH: 47
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 1..47
;   OTHER INFORMATION: polymorphic fragment 4-86-309
;   NAME/KEY: allele
;   LOCATION: 24
;   OTHER INFORMATION: polymorphic base A
;   NAME/KEY: primer_bind
;   LOCATION: 1..23
;   OTHER INFORMATION: potential microsequencing oligo 4-86-309.mis1
;   NAME/KEY: primer_bind
;   LOCATION: 25..47
;   OTHER INFORMATION: complement potential microsequencing oligo 4-86-309.mis2
US-09-853-526-225

  Query Match             63.0%;  Score 12.6;  DB 9;  Length 47;
  Best Local Similarity   78.9%;  Pred. No. 1.9e+03;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 TGCTCCCTGTTTCATCAAT 20
           ||||| ||| ||| ||| |
Db      44 TGCTCTCTGGTTCCTCACT 26


RESULT 10
US-09-853-526-302/c
; Sequence 302, Application US/09853526
; Patent No. US20020165345A1
; GENERAL INFORMATION:
;  APPLICANT: Cohen, Daniel
;  APPLICANT:  Blumenfeld, Marta
;  APPLICANT:  Ilya, Chumakov
;  APPLICANT:  Bougueleret, Lydie
;  TITLE OF INVENTION: PROSTATE CANCER GENE
;  FILE REFERENCE: GENSET.18CP1CP
;  CURRENT APPLICATION NUMBER: US/09/853,526
;  CURRENT FILING DATE:  2001-05-11
;  PRIOR APPLICATION NUMBER: 09/338,907
;  PRIOR FILING DATE: 1999-06-23
;  PRIOR APPLICATION NUMBER: 08/996,306
;  PRIOR FILING DATE: 1997-12-22
;  PRIOR APPLICATION NUMBER: 60/099,658
;  PRIOR FILING DATE: 1998-09-09
;  PRIOR APPLICATION NUMBER: 09/218,207
;  PRIOR FILING DATE: 1998-12-22
;  NUMBER OF SEQ ID NOS: 578
;  SOFTWARE: Patent.pm
; SEQ ID NO 302
;   LENGTH: 47
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: 1..47
;   OTHER INFORMATION: polymorphic fragment 4-86-309, variant version of SEQ ID225
;   NAME/KEY: allele
;   LOCATION: 24
;   OTHER INFORMATION: base T ; A in SEQ ID225
;   NAME/KEY: primer_bind
;   LOCATION: 1..23
;   OTHER INFORMATION: potential microsequencing oligo 4-86-309.mis1
;   NAME/KEY: primer_bind
;   LOCATION: 25..47
;   OTHER INFORMATION: complement potential microsequencing oligo 4-86-309.mis2
US-09-853-526-302

  Query Match             63.0%;  Score 12.6;  DB 9;  Length 47;
  Best Local Similarity   78.9%;  Pred. No. 1.9e+03;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 TGCTCCCTGTTTCATCAAT 20
           ||||| ||| ||| ||| |
Db      44 TGCTCTCTGGTTCCTCACT 26


RESULT 11
US-09-901-484A-225/c
; Sequence 225, Application US/09901484A
; Patent No. US20020119460A1
; GENERAL INFORMATION:
;  APPLICANT: Cohen, Daniel
;  APPLICANT:  Blumenfeld, Marta
;  APPLICANT:  Chumakov, Ilya
;  APPLICANT:  Bougueleret, Lydie
;  TITLE OF INVENTION: Prostate Cancer Gene
;  FILE REFERENCE: GEN-T111XC3D2
;  CURRENT APPLICATION NUMBER: US/09/901,484A
;  CURRENT FILING DATE:  2001-07-09
;  PRIOR APPLICATION NUMBER: US 08/996,306
;  PRIOR FILING DATE: 1997-12-22
;  PRIOR APPLICATION NUMBER: US 60/099,658
;  PRIOR FILING DATE: 1998-09-09
;  PRIOR APPLICATION NUMBER: US 09/218,207
;  PRIOR FILING DATE: 1998-12-22
;  PRIOR APPLICATION NUMBER: US 09/338,907
;  PRIOR FILING DATE: 1999-06-23
;  PRIOR APPLICATION NUMBER: US 09/853,526
;  PRIOR FILING DATE: 2001-05-11
;  NUMBER OF SEQ ID NOS: 578
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 225
;   LENGTH: 47
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: allele
;   LOCATION: (1)..(47)
;   OTHER INFORMATION: polymorphic fragment 4-86-309
;   NAME/KEY: allele
;   LOCATION: (24)..(24)
;   OTHER INFORMATION: polymorphic base A
;   NAME/KEY: primer_bind
;   LOCATION: (1)..(23)
;   OTHER INFORMATION: potential microsequencing oligo 4-86-309.mis1
;   NAME/KEY: primer_bind
;   LOCATION: (25)..(47)
;   OTHER INFORMATION: complement potential microsequencing oligo 4-86-309.mis2
US-09-901-484A-225

  Query Match             63.0%;  Score 12.6;  DB 10;  Length 47;
  Best Local Similarity   78.9%;  Pred. No. 1.9e+03;
```

```
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 TGCTCCCTGTTTCATCAAT 20
           ||||| ||| ||| ||| |
Db      44 TGCTCTCTGGTTCCTCACT 26


RESULT 12
US-09-901-484A-302/c
; Sequence 302, Application US/09901484A
; Patent No. US20020119460A1
; GENERAL INFORMATION:
;  APPLICANT: Cohen, Daniel
;  APPLICANT:  Blumenfeld, Marta
;  APPLICANT:  Chumakov, Ilya
;  APPLICANT:  Bougueleret, Lydie
;  TITLE OF INVENTION: Prostate Cancer Gene
;  FILE REFERENCE: GEN-T111XC3D2
;  CURRENT APPLICATION NUMBER: US/09/901,484A
;  CURRENT FILING DATE:  2001-07-09
;  PRIOR APPLICATION NUMBER: US 08/996,306
;  PRIOR FILING DATE: 1997-12-22
;  PRIOR APPLICATION NUMBER: US 60/099,658
;  PRIOR FILING DATE: 1998-09-09
;  PRIOR APPLICATION NUMBER: US 09/218,207
;  PRIOR FILING DATE: 1998-12-22
;  PRIOR APPLICATION NUMBER: US 09/338,907
;  PRIOR FILING DATE: 1999-06-23
;  PRIOR APPLICATION NUMBER: US 09/853,526
;  PRIOR FILING DATE: 2001-05-11
;  NUMBER OF SEQ ID NOS: 578
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 302
;  LENGTH: 47
;  TYPE: DNA
;  ORGANISM: Homo sapiens
;  FEATURE:
;  NAME/KEY: allele
;  LOCATION: (1)..(47)
;  OTHER INFORMATION: polymorphic fragment 4-86-309, variant version of SEQ ID 225
;  NAME/KEY: allele
;  LOCATION: (24)..(24)
;  OTHER INFORMATION: polymorphic base T; A is SEQ ID 225
;  NAME/KEY: primer_bind
;  LOCATION: (1)..(23)
;  OTHER INFORMATION: potential microsequencing oligo 4-86-309.mis1
;  NAME/KEY: primer_bind
;  LOCATION: (25)..(47)
;  OTHER INFORMATION: complement potential microsequencing oligo 4-86-309.mis2
US-09-901-484A-302

  Query Match             63.0%;  Score 12.6;  DB 10;  Length 47;
  Best Local Similarity   78.9%;  Pred. No. 1.9e+03;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 TGCTCCCTGTTTCATCAAT 20
           ||||| ||| ||| ||| |
Db      44 TGCTCTCTGGTTCCTCACT 26


RESULT 13
US-09-969-373-3369
; Sequence 3369, Application US/09969373
; Patent No. US20020133852A1
; GENERAL INFORMATION:
;  APPLICANT: Effertz, Roger J.
;  APPLICANT:  Hauge, Brian M.
;  TITLE OF INVENTION: Soybean SSRs and Methods of Genotyping
;  FILE REFERENCE: 38-10(52679)A
;  CURRENT APPLICATION NUMBER: US/09/969,373
;  CURRENT FILING DATE:  2001-10-02
;  PRIOR APPLICATION NUMBER: US 09/754,853
;  PRIOR FILING DATE: 2001-01-05
;  PRIOR APPLICATION NUMBER: US 09/760,427
;  PRIOR FILING DATE: 2001-01-13
;  PRIOR APPLICATION NUMBER: US 09/855,768
;  PRIOR FILING DATE: 2001-05-15
;  NUMBER OF SEQ ID NOS: 4593
; SEQ ID NO 3369
;   LENGTH: 18
;   TYPE: DNA
;   ORGANISM: Glycine max
US-09-969-373-3369

  Query Match             62.0%;  Score 12.4;  DB 10;  Length 18;
  Best Local Similarity   92.9%;  Pred. No. 2e+03;
  Matches   13;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       4 CTCCCTGTTTCATC 17
           |||||||||||  ||
Db       2 CTCCCTGTTTCCTC 15


RESULT 14
US-09-969-373-3370/c
; Sequence 3370, Application US/09969373
; Patent No. US20020133852A1
; GENERAL INFORMATION:
;  APPLICANT: Effertz, Roger J.
;  APPLICANT:  Hauge, Brian M.
;  TITLE OF INVENTION: Soybean SSRs and Methods of Genotyping
;  FILE REFERENCE: 38-10(52679)A
;  CURRENT APPLICATION NUMBER: US/09/969,373
;  CURRENT FILING DATE:  2001-10-02
;  PRIOR APPLICATION NUMBER: US 09/754,853
;  PRIOR FILING DATE: 2001-01-05
;  PRIOR APPLICATION NUMBER: US 09/760,427
;  PRIOR FILING DATE: 2001-01-13
;  PRIOR APPLICATION NUMBER: US 09/855,768
;  PRIOR FILING DATE: 2001-05-15
;  NUMBER OF SEQ ID NOS: 4593
; SEQ ID NO 3370
;   LENGTH: 18
;   TYPE: DNA
;   ORGANISM: Glycine max
US-09-969-373-3370

  Query Match             62.0%;  Score 12.4;  DB 10;  Length 18;
  Best Local Similarity   92.9%;  Pred. No. 2e+03;
  Matches   13;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       4 CTCCCTGTTTCATC 17
           |||||||||||  ||
Db      17 CTCCCTGTTTCCTC 4


RESULT 15
US-09-969-373-919/c
; Sequence 919, Application US/09969373
; Patent No. US20020133852A1
; GENERAL INFORMATION:
;  APPLICANT: Effertz, Roger J.
;  APPLICANT:  Hauge, Brian M.
;  TITLE OF INVENTION: Soybean SSRs and Methods of Genotyping
;  FILE REFERENCE: 38-10(52679)A
;  CURRENT APPLICATION NUMBER: US/09/969,373
;  CURRENT FILING DATE:  2001-10-02
;  PRIOR APPLICATION NUMBER: US 09/754,853
;  PRIOR FILING DATE: 2001-01-05
;  PRIOR APPLICATION NUMBER: US 09/760,427
;  PRIOR FILING DATE: 2001-01-13
;  PRIOR APPLICATION NUMBER: US 09/855,768
;  PRIOR FILING DATE: 2001-05-15
;  NUMBER OF SEQ ID NOS: 4593
; SEQ ID NO 919
;   LENGTH: 73
```

```
;   TYPE: DNA
;   ORGANISM: Glycine max
US-09-969-373-919

  Query Match            62.0%;  Score 12.4;  DB 10;  Length 73;
  Best Local Similarity  92.9%;  Pred. No. 2.5e+03;
  Matches   13;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       4 CTCCCTGTTTCATC 17
           ||||||||||| ||
Db      72 CTCCCTGTTTCCTC 59


Search completed: November 23, 2002, 06:42:14
Job time : 18.25 secs
```

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 25, 2002, 09:10:06 ; Search time 755.55 Seconds
(without alignments)
428.707 Million cell updates/sec

Title: US-09-296-264-18
Perfect score: 20
Sequence: 1 gtgctccctgtttcatcaat 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 357874

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	14.2	71.0	70	17	BH619677	BH619677 1007061F0
c	2	14	70.0	68	17	BH850637	BH850637 SALK_0716
	3	13.8	69.0	49	17	AZ783979	AZ783979 2M0026K08
	4	13.8	69.0	79	9	AA119008	AA119008 mp61b06.r
c	5	13.8	69.0	96	17	AZ858019	AZ858019 2M0163I05
c	6	13.4	67.0	40	9	AI159650	AI159650 ue98b01.x
c	7	13.4	67.0	62	12	BF026795	BF026795 601672091
c	8	13.4	67.0	71	17	AZ849964	AZ849964 2M0151E07
	9	13.4	67.0	75	9	AJ239761	AJ239761 AJ239761
	10	13.4	67.0	93	9	AL672516	AL672516 AL672516
	11	13.2	66.0	61	9	AI762131	AI762131 wh65b04.x
	12	13.2	66.0	64	17	CNS02YOA	AL219907 Tetraodon
c	13	13.2	66.0	71	17	BH218328	BH218328 1006078D1
	14	13.2	66.0	71	17	BH848957	BH848957 SALK_0690
c	15	13.2	66.0	75	9	AL640397	AL640397 AL640397
c	16	13.2	66.0	84	17	AZ412309	AZ412309 1M0185O08
c	17	13.2	66.0	96	17	TA102E11P	AL462358 T. brucei
c	18	13	65.0	100	14	BU033796	BU033796 QHJ2D23.y
c	19	12.8	64.0	58	14	H25101	H25101 y143f03.r1
c	20	12.8	64.0	62	10	BE261364	BE261364 601148494
c	21	12.8	64.0	65	17	AZ779503	AZ779503 2M0016E02
	22	12.8	64.0	83	9	AA671672	AA671672 v105d05.r
c	23	12.8	64.0	87	17	AZ338537	AZ338537 1M0069G12
	24	12.8	64.0	90	9	AA072789	AA072789 mm76f04.r
	25	12.8	64.0	90	14	W98578	W98578 mg14a08.r1
	26	12.8	64.0	92	10	AW711390	AW711390 f2b01ne.f
c	27	12.8	64.0	98	14	BQ087683	BQ087683 Ts363 The
	28	12.6	63.0	52	13	BM431788	BM431788 1Duo27E1.
c	29	12.6	63.0	54	9	AA595957	AA595957 nn66d05.s
c	30	12.6	63.0	57	17	TA151F02P	AL467333 T. brucei
	31	12.6	63.0	61	17	BH631025	BH631025 1007096B0
	32	12.6	63.0	62	17	AZ307027	AZ307027 1M0008A22
c	33	12.6	63.0	69	17	AZ782052	AZ782052 2M0021O20
c	34	12.6	63.0	69	17	BH863114	BH863114 SALK_0931
c	35	12.6	63.0	70	9	AA921376	AA921376 ak56e04.s
c	36	12.6	63.0	70	17	BH863077	BH863077 SALK_0930
	37	12.6	63.0	72	17	AZ837468	AZ837468 2M0132F17
	38	12.6	63.0	72	17	BH631091	BH631091 1007096G0
	39	12.6	63.0	75	17	AZ417458	AZ417458 1M0193G07
c	40	12.6	63.0	77	17	AZ845724	AZ845724 2M0145H07
c	41	12.6	63.0	79	9	AI194849	AI194849 ui57e10.x
	42	12.6	63.0	81	14	R19122	R19122 yg23d08.r1
	43	12.6	63.0	85	9	AI278351	AI278351 ql89b07.x
c	44	12.6	63.0	86	13	BM430083	BM430083 1Duo29D2.
	45	12.6	63.0	91	17	AZ308902	AZ308902 1M0012F05

ALIGNMENTS

RESULT 1
BH619677/c
LOCUS BH619677 70 bp DNA linear GSS 30-JAN-2002
DEFINITION 1007061F04.2EL_x1 1007 - RescueMu Grid H Zea mays genomic, DNA sequence.
ACCESSION BH619677
VERSION BH619677.1 GI:18430659
KEYWORDS GSS.
SOURCE Zea mays.
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 70)
AUTHORS Walbot,V.
TITLE Maize genomic sequences found using engineered RescueMu transposon
JOURNAL Unpublished (2001)
COMMENT Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Possible ligation site of ends cut by 2 different endonucleases.
Reverse complemented post-ligation sequence from source sequence.
Plate: 1007061 column: 18
Class: transposon-tagged.
FEATURES Location/Qualifiers

```
     source          1. .70
                     /organism="Zea mays"
                     /cultivar="mixed background W23/A188/B73"
                     /db_xref="taxon:4577"
                     /clone_lib="1007 - RescueMu Grid H"
                     /tissue_type="leaf"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /note="Organ: leaf; Vector: RescueMu (engineered from
                     pBlueScript backbone); Site_1: BamHI; Site_2: BglII;
                     RescueMu is a 4.9 kb, modified maize Mu transposon
                     designed to allow plasmid rescue from total genomic DNA.
                     Mu elements insert preferentially into transcription
                     units.  For more information on RescueMu, go to the  web
                     site 'www.zmdb.iastate.edu' and follow the links for
                     'RescueMu.'  Grid H was grown at Berkeley in 2001.  DNA
                     was extracted from leaf punches,  double digested using
                     BamHI and BglII, and ligated to form circular plasmids.
                     DH10B cells were transformed and then screened on LB
                     plates with ampicillin."
BASE COUNT       30 a     11 c     13 g     16 t
ORIGIN

  Query Match             71.0%;  Score 14.2;  DB 17;  Length 70;
  Best Local Similarity   84.2%;  Pred. No. 7.3e+03;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 TGCTCCCTGTTTCATCAAT 20
           |||||||| |||| | |||
Db      48 TGCTCCCTTTTTCTTTAAT 30


RESULT 2
BH850637/c
LOCUS       BH850637                 68 bp    DNA     linear   GSS 13-JUN-2002
DEFINITION  SALK_071623.16.90.x Arabidopsis thaliana TDNA insertion lines
            Arabidopsis thaliana genomic clone SALK_071623.16.90.x, DNA
            sequence.
ACCESSION   BH850637
VERSION     BH850637.1  GI:21421508
KEYWORDS    GSS.
SOURCE      thale cress.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 68)
  AUTHORS   Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R., Gadrinab
            ,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L., Shinn,P.
            , Zimmerman,J. and Ecker,J.R.
  TITLE     A Sequence-Indexed Library of Insertion Mutations in the
            Arabidopsis Genome
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Joseph R. Ecker
            Salk Institute Genomic Analysis Laboratory (SIGnAL)
            The Salk Institute for Biological Studies
            10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
            Tel: 858 453 4100 x1752
            Fax: 858 558 6379
            Email: ecker@salk.edu
            This is single pass sequence recovered from the left border of
            TDNA. This sequence lies within 300 bases of the 5' end of
            At5g02180.
            Class: TDNA tagged.
FEATURES             Location/Qualifiers
     source          1. .68
                     /organism="Arabidopsis thaliana"
                     /strain="Columbia 0"
                     /db_xref="taxon:3702"
                     /clone="SALK_071623.16.90.x"
                     /clone_lib="Arabidopsis thaliana TDNA insertion lines"
                     /note="PCR was performed on Arabidopsis thaliana lines
                     each of which contains  one or more TDNA insertion
                     elements. The resultant fragment for each  line was
                     directly sequenced to determine the genomic sequence at
                     the  site of insertion. Details of the protocols used can
                     be found at  http://signal.salk.edu/tdna_protocols.html"
BASE COUNT       18 a     12 c     18 g     20 t
ORIGIN

  Query Match             70.0%;  Score 14;  DB 17;  Length 68;
  Best Local Similarity   100.0%;  Pred. No. 9.1e+03;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       6 CCCTGTTTCATCAA 19
           ||||||||||||||
Db      51 CCCTGTTTCATCAA 38


RESULT 3
AZ783979
LOCUS       AZ783979                 49 bp    DNA     linear   GSS 16-FEB-2001
DEFINITION  2M0026K08F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC2M0026K08 F, DNA sequence.
ACCESSION   AZ783979
VERSION     AZ783979.1  GI:12919254
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 49)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0026  row: K  column: 08
            Seq primer: CGTTGTAAAACGACGGCCAGT
            Class: plasmid ends
            High quality sequence stop: 49.
FEATURES             Location/Qualifiers
     source          1. .49
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC2M0026K08"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
```

purified. The sheared, adaptored mouse DNA was annealed to adaptored vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."
BASE COUNT 9 a 14 c 7 g 19 t
ORIGIN

Query Match 69.0%; Score 13.8; DB 17; Length 49;
Best Local Similarity 88.2%; Pred. No. 1e+04;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 TGCTCCCTGTTTCATCA 18
||||| ||||| |||||
Db 3 TGCTCACTGTTCCATCA 19

RESULT 4
AA119008
LOCUS AA119008 79 bp mRNA linear EST 17-FEB-1997
DEFINITION mp61b06.r1 Soares_thymus_2NbMT Mus musculus cDNA clone IMAGE:573683 5', mRNA sequence.
ACCESSION AA119008
VERSION AA119008.1 GI:1676497
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 79)
AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.
TITLE The WashU-HHMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Marra M/Mouse EST Project
WashU-HHMI Mouse EST Project
Washington University School of MedicineP
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:348331
Seq primer: -28M13 rev2 from Amersham
High quality sequence stop: 61.
FEATURES Location/Qualifiers
source 1. .79
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:573683"
/clone_lib="Soares_thymus_2NbMT"
/sex="male"
/tissue_type="Thymus"
/dev_stage="4 weeks"
/lab_host="DH10B"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGGAGCGGCCGCGTTTTTTTTTTTTTTTTTTTTTT 3']; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. RNA provided by Dr. Bertrand Jordan. Library went through two rounds of normalization, and was constructed by Bento Soares and M.Fatima Bonaldo."
BASE COUNT 13 a 25 c 20 g 21 t
ORIGIN

Query Match 69.0%; Score 13.8; DB 9; Length 79;
Best Local Similarity 88.2%; Pred. No. 1.2e+04;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 TGCTCCCTGTTTCATCA 18
||||||||||| ||||
Db 28 TGCTCCCTGTTAAATCA 44

RESULT 5
AZ858019/c
LOCUS AZ858019 96 bp DNA linear GSS 21-FEB-2001
DEFINITION 2M0163I05F Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC2M0163I05 F, DNA sequence.
ACCESSION AZ858019
VERSION AZ858019.1 GI:13050748
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 96)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D.,Weiss,R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0163 row: I column: 05
Seq primer: CGTTGTAAAACGACGGCCAGT
Class: plasmid ends
High quality sequence stop: 96.
FEATURES Location/Qualifiers
source 1. .96
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC2M0163I05"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptored DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptored mouse DNA was annealed to adaptored vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."
BASE COUNT 26 a 21 c 32 g 17 t
ORIGIN

Query Match 69.0%; Score 13.8; DB 17; Length 96;

```
  Best Local Similarity   88.2%;  Pred. No. 1.3e+04;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       3 GCTCCCTGTTTCATCAA 19
           | ||||  ||||||||||
Db      18 GATCCCAGTTTCATCAA 2


RESULT 6
AI159650/c
LOCUS       AI159650                  40 bp    mRNA    linear   EST 02-OCT-1998
DEFINITION  ue98b01.x1 Sugano mouse embryo mewa Mus musculus cDNA clone
            IMAGE:1499113 3' similar to SW:PEX6_RAT P54777 PEROXISOME ASSEMBLY
            FACTOR-2 ;, mRNA sequence.
ACCESSION   AI159650
VERSION     AI159650.1  GI:3692832
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 40)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:936717
            Trace considered overall poor quality
            Possible reversed clone: similarity on wrong strand
            Seq primer: custom primer used
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .40
                     /organism="Mus musculus"
                     /strain="C57BL"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:1499113"
                     /clone_lib="Sugano mouse embryo mewa"
                     /dev_stage="embryo, 14 dpc"
                     /lab_host="DH10B"
                     /note="Vector: pME18S-FL3; Site_1: DraIII (CACTGTGTG);
                     Site_2: DraIII (CACCATGTG); 1st strand cDNA was primed
                     with an oligo(dT) primer  [ATGTGGCCTTTTTTTTTTTTTTTT];
                     double-stranded cDNA was ligated to a  DraIII adaptor
                     [TGTTGGCCTACTGG], digested and cloned into distinct DraIII
                     sites of the pME18S-FL3 vector (5' site CACTGTGTG, 3' site
                     CACCATGTG). XhoI should be used to isolate the cDNA
                     insert.  Size selection was performed to exclude fragments
                     <1.5kb.  Library constructed by Dr. Sumio Sugano
                     (University of Tokyo Institute of Medical Science).
                     Custom primers  for sequencing: 5' end primer
                     CTTCTGCTCTAAAAGCTGCG and 3' end primer
                     CGACCTGCAGCTCGAGCACA."
BASE COUNT       14 a     11 c      9 g      6 t
ORIGIN

  Query Match             67.0%;  Score 13.4;  DB 9;  Length 40;
  Best Local Similarity   93.3%;  Pred. No. 1.5e+04;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 TGCTCCCTGTTTCAT 16
           ||||||||||| ||||
Db      17 TGCTCCCTGTATCAT 3


RESULT 7
BF026795/c
LOCUS       BF026795                  62 bp    mRNA    linear   EST 10-OCT-2000
DEFINITION  601672091F1 NIH_MGC_20 Homo sapiens cDNA clone IMAGE:3954839 5',
            mRNA sequence.
ACCESSION   BF026795
VERSION     BF026795.1  GI:10734507
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 62)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC/DCTD/DTP
             cDNA Library Preparation: Ling Hong/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
            Plate: LLCM828  row: f  column: 24.
FEATURES             Location/Qualifiers
     source          1. .62
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3954839"
                     /clone_lib="NIH_MGC_20"
                     /tissue_type="melanotic melanoma"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: skin; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming.  Directionally
                     cloned into EcoRI/XhoI sites using the following 5'
                     adaptor: GGCACGAG(G).  Size-selected >500bp for average
                     insert size 1.8kb.  Library constructed by Ling Hong in
                     the laboratory of Gerald M. Rubin (University of
                     California, Berkeley) using  ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT       19 a     15 c      26 g      2 t
ORIGIN

  Query Match             67.0%;  Score 13.4;  DB 12;  Length 62;
  Best Local Similarity   93.3%;  Pred. No. 1.7e+04;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       4 CTCCCTGTTTCATCA 18
           |||||||||||  |||
Db      21 CTCCCTGTTTCTTCA 7


RESULT 8
AZ849964/c
LOCUS       AZ849964                  71 bp    DNA     linear   GSS 21-FEB-2001
DEFINITION  2M0151E07R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC2M0151E07 R, DNA sequence.
ACCESSION   AZ849964
VERSION     AZ849964.1  GI:13034499
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 71)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
```

```
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0151  row: E  column: 07
            Seq primer: CACACAGGAAACAGCTATGACC
            Class: plasmid ends
            High quality sequence stop: 71.
FEATURES             Location/Qualifiers
     source          1. .71
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC2M0151E07"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
BASE COUNT       13 a     14 c     21 g     23 t
ORIGIN

  Query Match             67.0%;  Score 13.4;  DB 17;  Length 71;
  Best Local Similarity   93.3%;  Pred. No. 1.8e+04;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       3 GCTCCCTGTTTCATC 17
           ||||||||| ||||
Db      39 GCTCCCTGTTACATC 25


RESULT 9
AJ239761
LOCUS       AJ239761                 75 bp    mRNA    linear   EST 10-AUG-1999
DEFINITION  AJ239761 Aspergillus niger ATCC6275 Aspergillus niger cDNA clone
            AN01F01, mRNA sequence.
ACCESSION   AJ239761
VERSION     AJ239761.1  GI:5443752
KEYWORDS    EST.
SOURCE      Aspergillus niger.
  ORGANISM  Aspergillus niger
            Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
            Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
REFERENCE   1  (bases 1 to 75)
  AUTHORS   Choi,J.Y., Lee,D.W., Koh,J.S., Kim,J.H., Yang,M.S. and Chae,K.S.
  TITLE     Identification of expressed sequence tags (ESTs) of the highly
            transcribed genes in Aspergillus niger
  JOURNAL   Biotechnol. Lett. 21, 381-384 (1999)
COMMENT     Contact: Chae KS
            Faculty of Biological Sciences
            Chonbuk National University
            Chonju 561-756, Republic of Korea.
FEATURES             Location/Qualifiers
     source          1. .75
                     /organism="Aspergillus niger"
                     /strain="ATCC6275"
                     /db_xref="taxon:5061"
                     /clone="AN01F01"
                     /clone_lib="Aspergillus niger ATCC6275"
BASE COUNT       27 a     12 c      9 g     22 t      5 others
ORIGIN

  Query Match             67.0%;  Score 13.4;  DB 9;  Length 75;
  Best Local Similarity   93.3%;  Pred. No. 1.9e+04;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       5 TCCCTGTTTCATCAA 19
           |||||||||| |||
Db      58 TCCCTGTTTCAACAA 72


RESULT 10
AL672516
LOCUS       AL672516                 93 bp    mRNA    linear   EST 18-MAR-2002
DEFINITION  AL672516 XGC-gastrula Silurana tropicalis cDNA clone TGas054e18 5',
            mRNA sequence.
ACCESSION   AL672516
VERSION     AL672516.1  GI:19528872
KEYWORDS    EST.
SOURCE      western clawed frog.
  ORGANISM  Silurana tropicalis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Silurana.
REFERENCE   1  (bases 1 to 93)
  AUTHORS   Taylor,R., Ashurst,J.L., Croning,M.D.R., Zorn,A.M. and Rogers,J.
  TITLE     Sanger Xenopus tropicalis EST project 2002
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Taylor R
            Sanger Centre
            Hinxton, Cambridgeshire, CB10 1SA, UK
            Email: trop@sanger.ac.uk
            Sanger Xenopus tropicalis EST project 2001
            TROPICALIS_SEQUENCE_ID: TGas054e18.p1cSP6
            Sequencing primer: P1CSP6
            This sequence is from a Xenopus Gene Collection (XGC) library
            constructed by Aaron M. Zorn.
FEATURES             Location/Qualifiers
     source          1. .93
                     /organism="Silurana tropicalis"
                     /db_xref="taxon:8364"
                     /clone="TGas054e18"
                     /clone_lib="XGC-gastrula"
                     /dev_stage="gastrula (stages 10.5-13 mixed)"
                     /lab_host="Escherichia coli XL1-blue"
                     /note="Vector: pCS107; Site_1: EcoRI; Site_2: NotI; cDNA
                     was oligo dT primed from 5ug of poly A+ RNA from stages
                     10-13 gastrulae.  EcoRI-NotI cut cDNA was then ligated
                     into pCS107 with EcoRI at the 5' end and NotI at the 3'
                     end."
BASE COUNT       25 a     19 c     22 g     27 t
ORIGIN

  Query Match             67.0%;  Score 13.4;  DB 9;  Length 93;
  Best Local Similarity   93.3%;  Pred. No. 2e+04;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       5 TCCCTGTTTCATCAA 19
           | ||||||||||||
Db      64 TGCCTGTTTCATCAA 78
```

```
RESULT 11
AI762131
LOCUS       AI762131                 61 bp    mRNA    linear   EST 24-JUN-1999
DEFINITION  wh65b04.x1 NCI_CGAP_Kid11 Homo sapiens cDNA clone IMAGE:2385583 3'
            similar to SW:NPH2_BOVIN Q28145 NEUREXOPHILIN 2 ;, mRNA sequence.
ACCESSION   AI762131
VERSION     AI762131.1  GI:5177798
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 61)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html

            Trace considered overall poor quality
            Seq primer: -40UP from Gibco
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .61
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2385583"
                     /clone_lib="NCI_CGAP_Kid11"
                     /lab_host="DH10B"
                     /note="Organ: kidney; Vector: pT7T3D-Pac (Pharmacia) with
                     a modified polylinker; Site_1: Not I; Site_2: Eco RI;
                     Plasmid DNA from the normalized library NCI_CGAP_Kid3 was
                     prepared, and ss circles were made in vitro. Following HAP
                     purification, this DNA was used as tracer in a subtractive
                     hybridization reaction. The driver was PCR-amplified cDNAs
                     from a pool of 5,000 clones made from the same library
                     (cloneIDs 1322376-1323911, 1456007-1456775, and
                     1500552-1502855). Subtraction by Bento Soares and M.
                     Fatima Bonaldo. "
BASE COUNT       10 a     14 c     20 g     17 t
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 9;  Length 61;
  Best Local Similarity   83.3%;  Pred. No. 2.2e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       3 GCTCCCTGTTTCATCAAT 20
           |||  ||||||| |||||
Db      18 GCTTGCTGTTTCCTCAAT 35


RESULT 12
CNS02YOA
LOCUS       CNS02YOA                 64 bp    DNA     linear   GSS 15-MAY-2000
DEFINITION  Tetraodon nigroviridis genome survey sequence T7 end of clone
            181C07 of library G from Tetraodon nigroviridis, genomic survey
            sequence.
ACCESSION   AL219907
VERSION     AL219907.1  GI:7878726
KEYWORDS    GSS; genome survey sequence.
SOURCE      Tetraodon nigroviridis.
  ORGANISM  Tetraodon nigroviridis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
            Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
            Tetraodontidae; Tetraodon.
REFERENCE   1  (bases 1 to 64)
  AUTHORS   Roest-Crollius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C.,
            Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F.,
            Saurin,W. and Weissenbach,J.
  TITLE     Human gene number estimate provided by genome wide analysis using
            Tetraodon nigroviridis DNA sequence
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 64)
  AUTHORS   Roest-Crollius,H., Jaillon,O., Dasilva,C., Fizames,C., Fisher,C.,
            Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and
            Weissenbach,J.
  TITLE     Charaterization and repeat analysis of the compact genome of the
            freshwater pufferfish Tetraodon nigroviridis
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 64)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-APR-2000)
COMMENT     This sequence is a single read and was generated as part of a large
            scale clone-end sequencing project of the Tetraodon nigroviridis
            genome. For more information, please take a look at
            http://www.genoscope.cns.fr/Tetraodon.
FEATURES             Location/Qualifiers
     source          1. .64
                     /organism="Tetraodon nigroviridis"
                     /db_xref="taxon:99883"
                     /clone="181C07"
                     /clone_lib="G"
                     /note="Genoscope sequence ID : C0AG181AB04LP1-end : T7"
BASE COUNT       1 a      17 c      12 g      33 t       1 others
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 17;  Length 64;
  Best Local Similarity   75.0%;  Pred. No. 2.2e+04;
  Matches   15;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 GTGCTCCCTGTTTCATCAAT 20
           ||||  |||||||| |: ||
Db      42 GTGCGGCCTGTTTCTTYCAT 61


RESULT 13
BH218328/c
LOCUS       BH218328                 71 bp    DNA     linear   GSS 08-NOV-2001
DEFINITION  1006078D10.2EL_x2 1006 - RescueMu Grid G Zea mays genomic, DNA
            sequence.
ACCESSION   BH218328
VERSION     BH218328.1  GI:16810986
KEYWORDS    GSS.
SOURCE      Zea mays.
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 71)
  AUTHORS   Walbot,V.
  TITLE     Maize genomic sequences found using engineered RescueMu transposon
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Walbot V
            Department of Biological Sciences
            Stanford University
            855 California Ave, Palo Alto, CA 94304, USA
            Tel: 650 723 2227
            Fax: 650 725 8221
            Email: walbot@stanford.edu
            Possible ligation site of ends cut by 2 different endonucleases.
            Reverse complemented post-ligation sequence from source sequence.
            Plate: 1006078  row: 27
            Class: transposon-tagged.
```

```
FEATURES             Location/Qualifiers
     source          1. .71
                     /organism="Zea mays"
                     /cultivar="mixed background W23/A188/B73"
                     /db_xref="taxon:4577"
                     /clone_lib="1006 - RescueMu Grid G"
                     /tissue_type="leaf"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /note="Organ: leaf; Vector: RescueMu (engineered from
                     pBlueScript backbone); Site_1: BamHI; Site_2: BglII;
                     RescueMu is a 4.9 kb, modified maize Mu transposon
                     designed to allow plasmid rescue from total genomic DNA.
                     Mu elements insert preferentially into transcription
                     units. For more information on RescueMu, go to the web
                     site 'www.zmdb.iastate.edu' and follow the links for
                     'RescueMu.' Grid G was grown at Stanford in 2000. DNA was
                     extracted from leaf punches, double digested using BamHI
                     and BglII, and ligated to form circular plasmids. DH10B
                     cells were transformed and then screened on LB plates with
                     ampicillin."
BASE COUNT       19 a     19 c     20 g     13 t
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 17;  Length 71;
  Best Local Similarity   83.3%;  Pred. No. 2.3e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 GTGCTCCCTGTTTCATCA 18
           |||| ||  |||||||||
Db      52 GTGCCCCGCGTTTCATCA 35


RESULT 14
BH848957
LOCUS       BH848957                 71 bp    DNA     linear   GSS 13-JUN-2002
DEFINITION  SALK_069069.23.60.x Arabidopsis thaliana TDNA insertion lines
            Arabidopsis thaliana genomic clone SALK_069069.23.60.x, DNA
            sequence.
ACCESSION   BH848957
VERSION     BH848957.1  GI:21419828
KEYWORDS    GSS.
SOURCE      thale cress.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 71)
  AUTHORS   Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R., Gadrinab
            ,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L., Shinn,P.
            , Zimmerman,J. and Ecker,J.R.
  TITLE     A Sequence-Indexed Library of Insertion Mutations in the
            Arabidopsis Genome
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Joseph R. Ecker
            Salk Institute Genomic Analysis Laboratory (SIGnAL)
            The Salk Institute for Biological Studies
            10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
            Tel: 858 453 4100 x1752
            Fax: 858 558 6379
            Email: ecker@salk.edu
            This is single pass sequence recovered from the left border of
            TDNA. This sequence lies within an annotated intron of At3g26590.
            Class: TDNA tagged.
FEATURES             Location/Qualifiers
     source          1. .71
                     /organism="Arabidopsis thaliana"
                     /strain="Columbia 0"
                     /db_xref="taxon:3702"
                     /clone="SALK_069069.23.60.x"
                     /clone_lib="Arabidopsis thaliana TDNA insertion lines"
                     /note="PCR was performed on Arabidopsis thaliana lines
                     each of which contains  one or more TDNA insertion
                     elements. The resultant fragment for each  line was
                     directly sequenced to determine the genomic sequence at
                     the  site of insertion. Details of the protocols used can
                     be found at  http://signal.salk.edu/tdna_protocols.html"
BASE COUNT       20 a     12 c     13 g     26 t
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 17;  Length 71;
  Best Local Similarity   83.3%;  Pred. No. 2.3e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 GTGCTCCCTGTTTCATCA 18
           || || | ||||||||||
Db      37 GTCCTTCTTGTTTCATCA 54


RESULT 15
AL640397/c
LOCUS       AL640397                 75 bp    mRNA    linear   EST 12-DEC-2001
DEFINITION  AL640397 XGC-neurula Silurana tropicalis cDNA clone TNeu002o02 5',
            mRNA sequence.
ACCESSION   AL640397
VERSION     AL640397.1  GI:16792528
KEYWORDS    EST.
SOURCE      western clawed frog.
  ORGANISM  Silurana tropicalis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Silurana.
REFERENCE   1  (bases 1 to 75)
  AUTHORS   Huckle,E., Taylor,R., Ashurst,J.L., Zorn,A.M. and Rogers,J.
  TITLE     Sanger Xenopus tropicalis EST project 2001 (10_2001)
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Huckle E
            Sanger Centre
            Hinxton, Cambridgeshire, CB10 1SA, UK
            Email: trop@sanger.ac.uk
            Sanger Xenopus tropicalis EST project 2001
            TROPICALIS_SEQUENCE_ID: TNeu002o02.sp6
            Sequencing primer: SP6
            This sequence is from a Xenopus Gene Collection (XGC) library
            constructed by Aaron M. Zorn.
FEATURES             Location/Qualifiers
     source          1. .75
                     /organism="Silurana tropicalis"
                     /db_xref="taxon:8364"
                     /clone="TNeu002o02"
                     /clone_lib="XGC-neurula"
                     /dev_stage="neurula"
                     /lab_host="Escherichia coli DH10B"
                     /note="Vector: pCS107; Site_1: EcoRI; Site_2: NotI; cDNA
                     was oligo dT primed from 5ug of poly A+ RNA from neurula.
                     EcoRI-NotI cut cDNA was then ligated into pCS107 with
                     EcoRI at the 5' end and NotI at the 3' end."
BASE COUNT       20 a     27 c     18 g      9 t      1 others
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 9;  Length 75;
  Best Local Similarity   83.3%;  Pred. No. 2.3e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       3 GCTCCCTGTTTCATCAAT 20
           | || ||||||| |||||
Db      46 GGTCTCTGTTTCTTCAAT 29


Search completed: November 26, 2002, 04:09:19
Job time : 766.8 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 3, 2002, 14:46:40 ; Search time 302.2 Seconds
(without alignments)
1926.063 Million cell updates/sec

Title: US-09-296-264-19
Perfect score: 20
Sequence: 1 cattgcctggcttcctggag 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 995600

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
c	1	15	75.0	66	6	A08800	A08800 Nucleotide
c	2	14.2	71.0	36	6	A51399	A51399 Sequence 1
	3	14.2	71.0	85	9	S65701	S65701 CRK=proto-o
c	4	13.8	69.0	59	6	AX011407	AX011407 Sequence
c	5	13.8	69.0	95	10	MMGXAD30	X75144 M.musculus
c	6	13.6	68.0	15	6	AX362584	AX362584 Sequence
c	7	13.6	68.0	40	6	AX040151	AX040151 Sequence
c	8	13.6	68.0	59	6	AX474322	AX474322 Sequence
	9	13.6	68.0	76	6	A25891	A25891 Artificial
	10	13.4	67.0	45	9	HUMKRT10CT	L20219 Human kerat
	11	13.4	67.0	74	6	I28815	I28815 Sequence 44
c	12	13.4	67.0	83	6	I28817	I28817 Sequence 46
	13	13.2	66.0	20	6	AR212060	AR212060 Sequence
c	14	13.2	66.0	24	6	AX290171	AX290171 Sequence
c	15	13.2	66.0	25	6	AX115768	AX115768 Sequence
	16	13.2	66.0	30	6	BD004772	BD004772 Novel VEG
c	17	13.2	66.0	50	6	AX162946	AX162946 Sequence
c	18	13.2	66.0	51	6	AX115769	AX115769 Sequence
c	19	13.2	66.0	51	6	AX162945	AX162945 Sequence
c	20	13.2	66.0	51	6	AX162947	AX162947 Sequence
c	21	13.2	66.0	51	6	AX162949	AX162949 Sequence
	22	13.2	66.0	51	6	AX204062	AX204062 Sequence
	23	13.2	66.0	65	6	AX484330	AX484330 Sequence
c	24	13.2	66.0	78	9	HSEEKD2	X59291 Human eek g
	25	13	65.0	17	6	A75919	A75919 Sequence 19
	26	13	65.0	17	6	AR048977	AR048977 Sequence
	27	13	65.0	17	6	AR064278	AR064278 Sequence
	28	13	65.0	17	6	AR164671	AR164671 Sequence
	29	13	65.0	39	6	AX038722	AX038722 Sequence
c	30	12.8	64.0	19	6	AR059493	AR059493 Sequence
c	31	12.8	64.0	19	6	E03000	E03000 DNA encodin
	32	12.8	64.0	19	6	E07058	E07058 Probe for H
c	33	12.8	64.0	19	6	E07082	E07082 Probe for H
c	34	12.8	64.0	19	6	I56677	I56677 Sequence 3
	35	12.8	64.0	21	8	ATH493632	AJ493632 Arabidops
	36	12.8	64.0	21	8	ATH493633	AJ493633 Arabidops
	37	12.8	64.0	21	8	ATH493634	AJ493634 Arabidops
	38	12.8	64.0	24	6	AR036150	AR036150 Sequence
c	39	12.8	64.0	27	6	AX117248	AX117248 Sequence
c	40	12.8	64.0	30	6	AR009009	AR009009 Sequence
c	41	12.8	64.0	36	6	AR169960	AR169960 Sequence
c	42	12.8	64.0	37	6	AR182323	AR182323 Sequence
c	43	12.8	64.0	37	6	E11725	E11725 Primer for
	44	12.8	64.0	51	6	AX117249	AX117249 Sequence
c	45	12.8	64.0	51	6	AX165481	AX165481 Sequence

ALIGNMENTS

RESULT 1
A08800/c
LOCUS A08800 66 bp DNA linear PAT 09-AUG-1993
DEFINITION Nucleotide sequence 1 from patent number WO8912457.
ACCESSION A08800
VERSION A08800.1 GI:411753
KEYWORDS .
SOURCE unidentified.
ORGANISM unidentified
unclassified.
FEATURES Location/Qualifiers
source 1. .66
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 20 a 16 c 17 g 13 t

```
ORIGIN

  Query Match            75.0%;  Score 15;  DB 6;  Length 66;
  Best Local Similarity  100.0%;  Pred. No. 6.2e+03;
  Matches   15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      6 CCTGGCTTCCTGGAG 20
          ||||||||||||||||
Db     43 CCTGGCTTCCTGGAG 29


RESULT 2
A51399/c
LOCUS       A51399                    36 bp    DNA     linear   PAT 10-MAR-1997
DEFINITION  Sequence 1 from Patent WO9617069.
ACCESSION   A51399
VERSION     A51399.1  GI:2304218
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 36)
  AUTHORS   Kossmann,J. and Springer,F.
  TITLE     TRANSGENIC PLANTS WITH IMPROVED BIOMASS PRODUCTION
  JOURNAL   Patent: WO 9617069-A 1 06-JUN-1996;
            INST GENBIOLOGISCHE FORSCHUNG (DE)
COMMENT     Other publication AU 4177096 960619
            Other publication DE 4444460 960530.
FEATURES             Location/Qualifiers
     source          1. .36
                     /organism="unidentified"
                     /db_xref="taxon:32644"
BASE COUNT       12 a      8 c      13 g      3 t
ORIGIN

  Query Match            71.0%;  Score 14.2;  DB 6;  Length 36;
  Best Local Similarity  84.2%;  Pred. No. 1.7e+04;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 CATTGCCTGGCTTCCTGGA 19
          ||| ||| || ||||||||
Db     24 CATGGCCCGGGTTCCTGGA 6


RESULT 3
S65701
LOCUS       S65701                    85 bp    DNA     linear   PRI 30-NOV-1993
DEFINITION  CRK=proto-oncogene [intron/exon junction] [human, Genomic, 85 nt].
ACCESSION   S65701
VERSION     S65701.1  GI:430792
KEYWORDS    .
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 85)
  AUTHORS   Fioretos,T., Heisterkamp,N., Groffen,J., Benjes,S. and Morris,C.
  TITLE     CRK proto-oncogene maps to human chromosome band 17p13
  JOURNAL   Oncogene 8 (10), 2853-2855 (1993)
  MEDLINE   93390962
   PUBMED   8378094
  REMARK    GenBank staff at the National Library of Medicine created this
            entry [NCBI gibbsq 137869] from the original journal article.
            This sequence comes from Fig. 1.
            Map location: 17p13.
FEATURES             Location/Qualifiers
     source          1. .85
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     gene            14. .85
                     /partial
                     /gene="CRK"
                     /note="proto-oncogene"
     CDS             14. .85
                     /partial
                     /gene="CRK"
                     /note="proto-oncogene; This sequence comes from Fig. 1"
                     /codon_start=1
                     /protein_id="AAB28213.1"
                     /db_xref="GI:430793"
                     /translation="VSPSRLRIGDQEFDSLPALLEFYK"
BASE COUNT       21 a      19 c      19 g      26 t
ORIGIN

  Query Match            71.0%;  Score 14.2;  DB 9;  Length 85;
  Best Local Similarity  84.2%;  Pred. No. 1.6e+04;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 CATTGCCTGGCTTCCTGGA 19
          |||||||||  || |||||
Db     57 CATTGCCTGCTTTACTGGA 75


RESULT 4
AX011407/c
LOCUS       AX011407                  59 bp    DNA     linear   PAT 06-SEP-2000
DEFINITION  Sequence 84 from Patent WO9955907.
ACCESSION   AX011407
VERSION     AX011407.1  GI:9997957
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 59)
  AUTHORS   Koetter,P., Entian,K.D. and Diu-Hercend,A.
  TITLE     Method for screening antimycotic substances using essential genes
            from s. Cerevisiae
  JOURNAL   Patent: WO 9955907-A 84 04-NOV-1999;
            KOETTER PETER (DE); ENTIAN KARL DIETER (DE); DIU HERCEND ANITA
            (FR); HOECHST MARION ROUSSEL INC (FR)
FEATURES             Location/Qualifiers
     source          1. .59
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="primer YML127w-S1"
BASE COUNT       16 a      20 c      14 g      9 t
ORIGIN

  Query Match            69.0%;  Score 13.8;  DB 6;  Length 59;
  Best Local Similarity  88.2%;  Pred. No. 2.6e+04;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      4 TGCCTGGCTTCCTGGAG 20
          ||||| |||| ||||||
Db     30 TGCCTCGCTTACTGGAG 14


RESULT 5
MMGXAD30/c
LOCUS       MMGXAD30                  95 bp    DNA     linear   ROD 27-JUL-1995
DEFINITION  M.musculus (129/Sv) gene for xanthine dehydrogenase, exon 30.
ACCESSION   X75144
VERSION     X75144.1  GI:473064
KEYWORDS    xanthine dehydrogenase; XD gene.
SOURCE      Mus musculus.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 95)
  AUTHORS   Cazzaniga,G.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-SEP-1993) G. Cazzaniga, Istituto di Ricerche
            Farmacologiche, 'Mario Negri', Via Eritrea 62, 20157 Milano, ITALY
REFERENCE   2  (bases 1 to 95)
  AUTHORS   Cazzaniga,G., Terao,M., Lo Schiavo,P., Galbiati,F., Segalla,F.,
            Seldin,M.F. and Garattini,E.
```

```
  TITLE     Chromosomal mapping, isolation, and characterization of the mouse
            xanthine dehydrogenase gene
  JOURNAL   Genomics 23 (2), 390-402 (1994)
  MEDLINE   95137585
   PUBMED   7835888
FEATURES             Location/Qualifiers
     source          1. .95
                     /organism="Mus musculus"
                     /strain="129/Sv"
                     /db_xref="taxon:10090"
                     /clone="4"
                     /tissue_type="spleen"
                     /clone_lib="mXD15"
     gene            11. .85
                     /gene="XD"
     exon            11. .85
                     /gene="XD"
                     /number=30
                     /usedin=X75129:XDCDS
                     /usedin=X75129:XDMRNA
                     /label=ex30
BASE COUNT       28 a     24 c     25 g     18 t
ORIGIN

  Query Match             69.0%;  Score 13.8;  DB 10;  Length 95;
  Best Local Similarity   88.2%;  Pred. No. 2.6e+04;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      4 TGCCTGGCTTCCTGGAG 20
          || | ||||||||||||
Db     21 TGACAGGCTTCCTGGAG 5


RESULT 6
AX362584/c
LOCUS       AX362584                 15 bp    DNA     linear   PAT 15-FEB-2002
DEFINITION  Sequence 18 from Patent WO0208425.
ACCESSION   AX362584
VERSION     AX362584.1  GI:18694728
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Finkel,K. and Koshy,B.
  TITLE     Haplotypes of the adrb3 gene
  JOURNAL   Patent: WO 0208425-A 18 31-JAN-2002;
            Genaissance Pharmaceuticals, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .15
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT        3 a      5 c      5 g      1 t      1 others
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 6;  Length 15;
  Best Local Similarity   92.9%;  Pred. No. 3.5e+04;
  Matches   13;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      7 CTGGCTTCCTGGAG 20
          |:|||||||||||
Db     15 CYGGCTTCCTGGAG 2


RESULT 7
AX040151/c
LOCUS       AX040151                 40 bp    DNA     linear   PAT 18-NOV-2000
DEFINITION  Sequence 47 from Patent WO0063438.
ACCESSION   AX040151
VERSION     AX040151.1  GI:11230101
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 40)
  AUTHORS   Gould-Rothberg,B.E. and Rastelli,L.
  TITLE     Method of classifying a thyroid carcinoma using differential gene
            expression
  JOURNAL   Patent: WO 0063438-A 47 26-OCT-2000;
            Curagen Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .40
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     variation       18
                     /note="FOUND TO BE G IN 2/32 IN SAMPLED POPULATION"
BASE COUNT       10 a     13 c      9 g      8 t
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 6;  Length 40;
  Best Local Similarity   80.0%;  Pred. No. 3.4e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 CATTGCCTGGCTTCCTGGAG 20
          | || ||||||||| | |||
Db     20 CCTTTCCTGGCTTCTTAGAG 1


RESULT 8
AX474322/c
LOCUS       AX474322                 59 bp    DNA     linear   PAT 12-AUG-2002
DEFINITION  Sequence 83 from Patent EP1223218.
ACCESSION   AX474322
VERSION     AX474322.1  GI:22213916
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Fraser,C.C.
  TITLE     Cd2000 and cd2001 molecules and uses thereof
  JOURNAL   Patent: EP 1223218-A 83 17-JUL-2002;
            Millennium Pharmaceuticals, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .59
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT        8 a     17 c     17 g     17 t
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 6;  Length 59;
  Best Local Similarity   80.0%;  Pred. No. 3.4e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 CATTGCCTGGCTTCCTGGAG 20
          | ||| ||| | ||||||||
Db     56 CTTTGGCTGCCCTCCTGGAG 37


RESULT 9
A25891
LOCUS       A25891                   76 bp    DNA     linear   PAT 02-OCT-1995
DEFINITION  Artificial VLP oligonucleotide TB5A.
ACCESSION   A25891
VERSION     A25891.1  GI:1248137
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
FEATURES             Location/Qualifiers
     source          1. .76
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
```

```
BASE COUNT       25 a     14 c     23 g     14 t
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 6;  Length 76;
  Best Local Similarity   80.0%;  Pred. No. 3.3e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 CATTGCCTGGCTTCCTGGAG 20
          ||| | | ||||||| ||||
Db     55 CATGGGCAGGCTTCCAGGAG 74


RESULT 10
HUMKRT10CT
LOCUS       HUMKRT10CT               45 bp    DNA     linear   PRI 06-JAN-1995
DEFINITION  Human keratin 10 (KRT10) gene, partial cds including polymorphism.
ACCESSION   L20219
VERSION     L20219.1  GI:307089
KEYWORDS    keratin; keratin 10; polymorphism; suprabasal keratin.
SOURCE      Homo sapiens Adult Blood DNA.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 45)
  AUTHORS   Rothnagel,J.J., Dominey,A., Fisher,M., Axtell,S., Pittelkow,M.,
            Anton-Lamprecht,I., Hohl,D. and Roop,D.
  TITLE     Identification of mutational hot spots in the suprabasal keratin
            genes from patients with epidermolytic hyperkeratosis
  JOURNAL   Unpublished (1993)
FEATURES             Location/Qualifiers
     source          1. .45
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /map="17q21-q23"
                     /cell_type="Lymphocyte"
                     /tissue_type="Blood"
                     /dev_stage="Adult"
                     /note="from EHK patient"
     gene            1. .45
                     /gene="KRT10"
     CDS             <1. .>45
                     /gene="KRT10"
                     /function="intermediate filament precursor"
                     /note="helix-initiation motif of 1A segment of rod domain"
                     /codon_start=1
                     /product="keratin 10"
                     /protein_id="AAB59439.1"
                     /db_xref="GI:307090"
                     /db_xref="GDB:G00-118-828"
                     /translation="KVTMQNLNDCLASYL"
     variation       28
                     /gene="KRT10"
                     /note="polymorphism results in an Arg->Cys mutation at
                     amino acid position #10 in the HK10 rod; G00-118-828"
                     /phenotype="wild-type (C)"
                     /replace="c"
BASE COUNT       13 a     11 c      9 g     12 t
ORIGIN

  Query Match             67.0%;  Score 13.4;  DB 9;  Length 45;
  Best Local Similarity   93.3%;  Pred. No. 4.3e+04;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      2 ATTGCCTGGCTTCCT 16
          | ||||||||||||||
Db     26 ACTGCCTGGCTTCCT 40


RESULT 11
I28815
LOCUS       I28815                   74 bp    DNA     linear   PAT 06-FEB-1997
DEFINITION  Sequence 44 from patent US 5574007.
ACCESSION   I28815
VERSION     I28815.1  GI:1819591
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 74)
  AUTHORS   Zushi,M., Gomi,K., Yamamoto,S., Suzuki,K. and Matsuda,A.
  TITLE     Polypeptide capable of interacting with thrombin
  JOURNAL   Patent: US 5574007-A 44 12-NOV-1996;
FEATURES             Location/Qualifiers
     source          1. .74
                     /organism="unknown"
BASE COUNT        9 a     25 c     23 g     17 t
ORIGIN

  Query Match             67.0%;  Score 13.4;  DB 6;  Length 74;
  Best Local Similarity   93.3%;  Pred. No. 4.2e+04;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      4 TGCCTGGCTTCCTGG 18
          ||||||||| |||||
Db     43 TGCCTGGCTGCCTGG 57


RESULT 12
I28817/c
LOCUS       I28817                   83 bp    DNA     linear   PAT 06-FEB-1997
DEFINITION  Sequence 46 from patent US 5574007.
ACCESSION   I28817
VERSION     I28817.1  GI:1819593
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 83)
  AUTHORS   Zushi,M., Gomi,K., Yamamoto,S., Suzuki,K. and Matsuda,A.
  TITLE     Polypeptide capable of interacting with thrombin
  JOURNAL   Patent: US 5574007-A 46 12-NOV-1996;
FEATURES             Location/Qualifiers
     source          1. .83
                     /organism="unknown"
BASE COUNT       20 a     25 c     28 g     10 t
ORIGIN

  Query Match             67.0%;  Score 13.4;  DB 6;  Length 83;
  Best Local Similarity   93.3%;  Pred. No. 4.2e+04;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      4 TGCCTGGCTTCCTGG 18
          ||||||||| |||||
Db     45 TGCCTGGCTGCCTGG 31


RESULT 13
AR212060
LOCUS       AR212060                 20 bp    DNA     linear   PAT 20-JUN-2002
DEFINITION  Sequence 27 from patent US 6399379.
ACCESSION   AR212060
VERSION     AR212060.1  GI:21515546
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 20)
  AUTHORS   Baker,B.F. and Freier,S.M.
  TITLE     Antisense modulation of interleukin 12 p35 subunit expression
  JOURNAL   Patent: US 6399379-A 27 04-JUN-2002;
FEATURES             Location/Qualifiers
     source          1. .20
                     /organism="unknown"
BASE COUNT        1 a      5 c      6 g      8 t
ORIGIN
```

```
  Query Match             66.0%;  Score 13.2;  DB 6;  Length 20;
  Best Local Similarity   83.3%;  Pred. No. 5.7e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       3 TTGCCTGGCTTCCTGGAG 20
           ||| ||||| | ||||||
Db       2 TTGTCTGGCCTTCTGGAG 19


RESULT 14
AX290171/c
LOCUS       AX290171                 24 bp    DNA     linear   PAT 21-NOV-2001
DEFINITION  Sequence 1933 from Patent WO0179548.
ACCESSION   AX290171
VERSION     AX290171.1  GI:17051854
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Barany,F., Zirvi,M., Gerry,N.P., Favis,R. and Kliman,R.
  TITLE     Method of designing addressable array for detection of nucleic acid
            sequence differences using ligase detection reaction
  JOURNAL   Patent: WO 0179548-A 1933 25-OCT-2001;
            CORNELL RESEARCH FOUNDATION, INC. (US)
FEATURES             Location/Qualifiers
     source          1. .24
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="Hypothetical Probe Sequence"
BASE COUNT        8 a      6 c      7 g      3 t
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 6;  Length 24;
  Best Local Similarity   83.3%;  Pred. No. 5.6e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 ATTGCCTGGCTTCCTGGA 19
           |||| | |||||||| ||
Db      23 ATTGTCCGGCTTCCTCGA 6


RESULT 15
AX115768/c
LOCUS       AX115768                 25 bp    DNA     linear   PAT 11-MAY-2001
DEFINITION  Sequence 891 from Patent WO0129262.
ACCESSION   AX115768
VERSION     AX115768.1  GI:14032710
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 25)
  AUTHORS   Picoult-Newburg,L. and Pohl,M.
  TITLE     Genotyping reagents, kits and methods of use thereof
  JOURNAL   Patent: WO 0129262-A 891 26-APR-2001;
            Orchid BioSciences, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .25
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="Primer"
BASE COUNT        8 a      7 c      5 g      5 t
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 6;  Length 25;
  Best Local Similarity   83.3%;  Pred. No. 5.6e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 ATTGCCTGGCTTCCTGGA 19
           | || || ||||||||||
Db      18 AGTGGCTTGCTTCCTGGA 1


Search completed: December  3, 2002, 18:14:07
Job time : 309.2 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 05:52:31 ; Search time 98.55 Seconds
(without alignments)
457.027 Million cell updates/sec

Title: US-09-296-264-19
Perfect score: 20
Sequence: 1 cattgcctggcttcctggag 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 2390332

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	20	100.0	20	21	AAZ31449	Human neuropilin m
	2	15.4	77.0	59	20	AAX22821	RGC p53 binding si
c	3	15.2	76.0	24	24	ABL41245	Human neuregulin 5
c	4	15.2	76.0	79	13	AAQ33658	Sequence downstrea
c	5	14.8	74.0	60	24	ABN35160	Human spliced tran
	6	14.8	74.0	87	21	AAC14135	Human secreted pro
c	7	14.4	72.0	26	21	AAZ50174	PCR primer DAE6 fo
c	8	14.4	72.0	51	22	AAL29599	Human SNP oligonuc
c	9	14.4	72.0	65	24	ABN56037	Mouse spliced tran
c	10	14.2	71.0	36	17	AAT32566	E.coli polyphospha
	11	14	70.0	21	24	ABK65538	Human single nucle
	12	13.8	69.0	24	22	AAH26781	Mouse T cell recep
	13	13.8	69.0	36	21	AAA35734	Permutein linker e
c	14	13.8	69.0	59	21	AAZ96894	S. cerevisiae gene
	15	13.8	69.0	60	21	AAC60378	DNA encoding pepti
c	16	13.8	69.0	60	24	ABN32687	Human spliced tran
	17	13.8	69.0	60	24	ABL56810	Leader peptide seq
c	18	13.8	69.0	65	24	ABN53604	Mouse spliced tran
	19	13.8	69.0	71	18	AAV75958	Staphylococcus aur
c	20	13.6	68.0	15	24	ABK11467	ASO primer #3, use
c	21	13.6	68.0	51	19	AAV07047	Truncated barbouri
c	22	13.6	68.0	51	19	AAV07042	Truncated barbouri
c	23	13.6	68.0	51	19	AAV06924	Truncated barbouri
	24	13.6	68.0	65	24	ABN30477	Rat spliced transc
	25	13.6	68.0	76	14	AAQ36748	Sequence of oligo
	26	13.4	67.0	15	22	AAF52898	IGF-I oligonucleot
	27	13.4	67.0	15	22	AAF52899	IGF-I oligonucleot
	28	13.4	67.0	17	20	AAV93384	Human B-raf substr
	29	13.4	67.0	17	20	AAV93385	Human B-raf substr
	30	13.4	67.0	26	20	AAX04596	PCR primer L1Ns1 u
	31	13.4	67.0	32	18	AAT63580	PCR primer used in
	32	13.4	67.0	41	24	ABL58700	Polyadenylate bind
	33	13.4	67.0	41	24	ABN83990	Polyadenylate bind
	34	13.4	67.0	50	22	AAL31932	Human SNP oligonuc
c	35	13.4	67.0	65	24	ABN27736	Rat spliced transc
c	36	13.2	66.0	24	18	AAT51157	Homeoprotein regul
c	37	13.2	66.0	24	22	AAF56502	Arabidopsis CDC27B
c	38	13.2	66.0	24	24	ABI86262	Capture oligonucle
	39	13.2	66.0	24	24	ABI86263	Capture oligonucle
c	40	13.2	66.0	25	22	AAH38095	SNP specific SNPE
	41	13.2	66.0	30	22	AAF82274	Human PIGF cDNA PC
	42	13.2	66.0	30	24	ABN88989	Human PDGF PCR pri
	43	13.2	66.0	31	24	ABN87270	Lolium perenne LpO
c	44	13.2	66.0	40	21	AAC68533	Peflin SNP probe S
c	45	13.2	66.0	50	22	AAL29386	Human SNP oligonuc

ALIGNMENTS

RESULT 1
AAZ31449
ID AAZ31449 standard; DNA; 20 BP.
XX
AC AAZ31449;
XX
DT 07-FEB-2000 (first entry)
XX
DE Human neuropilin mRNA specific antisense oligo GTI3620.
XX
KW Neuropilin; human; growth; metastasis; tumor; neovascularisation;
KW cancer; papilloma; diabetic retinopathy; antisense; ss.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9955855-A2.
XX
PD 04-NOV-1999.
XX
PF 23-APR-1999; 99WO-CA00324.
XX
PR 23-APR-1998; 98US-0082791.
XX
PA (GENE-) GENESENSE TECHNOLOGIES INC.
XX
PI Wright JA, Young AH, Lee YS;
XX
DR WPI; 2000-023357/02.
XX
PT Antisense oligonucleotides that inhibit neuropilin expression, useful
PT for treating cancer -

```
XX
PS   Claim 4; Page 16; 57pp; English.
XX
CC   Sequences AAZ31431-460 represent antisense oligonucleotides which
CC   inhibit human neuropilin expression. The antisense oligonucleotides can
CC   be used to inhibit the growth or metastasis of a mammalian tumor and
CC   inhibit neovascularisation. The oligonucleotides may be used to treat
CC   various forms of cancers or tumors, such as sarcomas, melanomas,
CC   adenomas, carcinomas of solid tissue, hypoxic tumors, squamous cell
CC   carcinomas of the mouth, throat, larynx and lung, genitourinary cancers
CC   such as cervical and bladder cancer, hematopoietic cancers, colon cancer,
CC   breast cancer, pancreatic cancer, renal cancer, brain cancer, skin
CC   cancer, liver cancer, head and neck cancers, and nervous system cancers,
CC   as well as benign lesions such as papillomas. The methods may be used to
CC   treat neovascularisation disorders such as diabetic retinopathy, and
CC   retinopathy of prematurity and age related macular degeneration.
XX
SQ   Sequence 20 BP; 2 A; 6 C; 6 G; 6 T; 0 other;

  Query Match             100.0%;  Score 20;  DB 21;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 5.1;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CATTGCCTGGCTTCCTGGAG 20
           ||||||||||||||||||||
Db       1 CATTGCCTGGCTTCCTGGAG 20


RESULT 2
AAX22821
ID   AAX22821 standard; DNA; 59 BP.
XX
AC   AAX22821;
XX
DT   27-MAY-1999  (first entry)
XX
DE   RGC p53 binding site structure 3 DNA.
XX
KW   Modulator; p53; detection; binding; regulation; treatment; tumour;
KW   conformation; primer; ss.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   misc_binding    14..15
FT                   /*tag=  a
FT                   /note= "Binds to nucleotides 46..47"
FT   misc_binding    17..18
FT                   /*tag=  b
FT                   /note= "Binds to nucleotides 43..44"
FT   misc_binding    24..25
FT                   /*tag=  c
FT                   /note= "Binds to nucleotides 36..37"
FT   misc_binding    27..29
FT                   /*tag=  d
FT                   /note= "Binds to nucleotides 32..34"
FT   misc_binding    32..34
FT                   /*tag=  e
FT                   /note= "Binds to nucleotides 27..29"
FT   misc_binding    36..37
FT                   /*tag=  f
FT                   /note= "Binds to nucleotides 24..25"
FT   misc_binding    43..44
FT                   /*tag=  g
FT                   /note= "Binds to nucleotides 17..18"
FT   misc_binding    46..47
FT                   /*tag=  h
FT                   /note= "Binds to nucleotides 46..47"
XX
PN   WO9908712-A2.
XX
PD   25-FEB-1999.
XX
PF   07-AUG-1998;   98WO-DE02326.
XX
PR   14-AUG-1997;   97DE-1035221.
XX
PA   (DEPP/) DEPPERT W W.
XX
PI   Deppert WW,  Kim E;
XX
DR   WPI; 1999-180763/15.
XX
PT   Modulation of p53 binding to target genes - by modifying
PT   conformation of p53 or target gene
XX
PS   Example; Fig 1; 15pp; German.
XX
CC   This invention describes a method for modulating the binding of p53 to
CC   a target gene, in which the conformations of p53 and the target gene
CC   are tuned to one another and the binding of p53 is detected directly
CC   or indirectly. The products of this invention can be used to correct
CC   impaired binding of p53 or to alter the binding specificity of p53 so
CC   that specific target genes are regulated, especially in the treatment
CC   of tumours or to identify substances that act as conformation modulators
CC   for p53 and/or its target genes.
XX
SQ   Sequence 59 BP; 7 A; 17 C; 21 G; 14 T; 0 other;

  Query Match             77.0%;  Score 15.4;  DB 20;  Length 59;
  Best Local Similarity   94.1%;  Pred. No. 7.6e+02;
  Matches   16;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       3 TTGCCTGGCTTCCTGGA 19
           |||||||||| ||||||
Db      21 TTGCCTGGCTGCCTGGA 37


RESULT 3
ABL41245/c
ID   ABL41245 standard; DNA; 24 BP.
XX
AC   ABL41245;
XX
DT   16-MAY-2002  (first entry)
XX
DE   Human neuregulin 55 PCR primer SEQ ID NO 3.
XX
KW   Human; neuregulin 55; nervous system; development; neuropsychopathy;
KW   tumour; inflammation; immunological disease; primer; ss.
XX
OS   Homo sapiens.
XX
PN   CN1324826-A.
XX
PD   05-DEC-2001.
XX
PF   19-MAY-2000; 2000CN-0115761.
XX
PR   19-MAY-2000; 2000CN-0115761.
XX
PA   (BODE-) BODE GENE DEV CO LTD SHANGHAI.
XX
PI   Mao Y,  Xie Y;
XX
DR   WPI; 2002-217507/28.
XX
PT   New polypeptide human neuregulin 55 and polynucleotides for encoding
PT   same -
XX
PS   Example 3; Page 18 (Disclosure); 35pp; Chinese.
XX
CC   The invention relates to human neuregulin 55, polynucleotide for coding
CC   this polypeptide and a method for producing this polypeptide by using DNA
CC   recombination technique. The invention also discloses the method for
CC   curing several diseases, such as nervous system developmental diseases,
```

CC neuropsychopathy, other nervous system diseases, development disturbance,
CC tumours, inflammations and immunological disease by using said
CC polypeptide. The invention also discloses an antagonist for resisting
CC said polypeptide and its therapeutic action and also discloses the
CC application of polynucleotide to coding this novel human neuregulin 55.
CC The present sequence is that of a human neuregulin 55 primer, useful to
CC the invention.
XX
SQ Sequence 24 BP; 7 A; 9 C; 4 G; 4 T; 0 other;

Query Match 76.0%; Score 15.2; DB 24; Length 24;
Best Local Similarity 85.0%; Pred. No. 8.7e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CATTGCCTGGCTTCCTGGAG 20
||||||||||| | | |||||
Db 22 CATTGCCTGGGTACTTGGAG 3

RESULT 4
AAQ33658/c
ID AAQ33658 standard; DNA; 79 BP.
XX
AC AAQ33658;
XX
DT 02-FEB-1993 (first entry)
XX
DE Sequence downstream of a microsatellite from clone TGLA102.
XX
KW PCR; selection; primers; OPTIPRIM; breeding; cattle; parentage;
KW genetic mapping; traits; amplification; ss.
XX
OS Bos taurus.
XX
PN WO9213102-A.
XX
PD 06-AUG-1992.
XX
PF 15-JAN-1992; 92WO-US00340.
XX
PR 15-JAN-1991; 91US-0642342.
XX
PA (GENM-) GENMARK.
XX
PI Georges M, Massey JM;
XX
DR WPI; 1992-284684/34.
XX
PT Polymorphic bovine DNA markers - used in genetic identification,
PT gene mapping, and selective breeding
XX
PS Table 7; Page 193; 517pp; English.
XX
CC The sequence is from downstream of a bovine microsatellite sequence
CC obtd. by screening a library of bovine MboI DNA fragments of between
CC 250 and 500 bp with an (AC)15 and a (TC)15 oligonucleotide probe.
CC One out of 50 clones cross-hybridised. Assuming independent
CC distribution of microsatellites and MboI sites, the frequency of
CC (T6)n >9 microsatellites in the bovine genome is estimated at >100,
CC 000. The sequence information for ca. 230 such bovine microsatellites
CC is summarised in the specification and indexed herein (see below).
CC The sequences upstream and downstream of the microsatellite sequence
CC were used to generate the required PCR primers for in vitro
CC amplification of the corresp. microsatellite (using the program
CC OPTIPRIM). The microsatellites may be used to identify individuals,
CC for parentage testing, and in the genetic mapping of economic trait
CC loci, or genes involved the determinism of economically important
CC traits esp. in cattle, to allow selective breeding.
CC See also AAQ33501-34437.
XX
SQ Sequence 79 BP; 24 A; 20 C; 21 G; 14 T; 0 other;

Query Match 76.0%; Score 15.2; DB 13; Length 79;
Best Local Similarity 85.0%; Pred. No. 9.7e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CATTGCCTGGCTTCCTGGAG 20
| ||||||| |||| |||||
Db 70 CCTTGCCTGCCTTCTTGGAG 51

RESULT 5
ABN35160/c
ID ABN35160 standard; DNA; 60 BP.
XX
AC ABN35160;
XX
DT 15-JUL-2002 (first entry)
XX
DE Human spliced transcript detection oligonucleotide SEQ ID NO:7908.
XX
KW Human; mouse; rat; splice transcript; detection; RNA transcript;
KW splice variant; transcriptome; oligonucleotide library; ss.
XX
OS Homo sapiens.
XX
PN WO200210449-A2.
XX
PD 07-FEB-2002.
XX
PF 20-JUL-2001; 2001WO-IB01903.
XX
PR 28-JUL-2000; 2000US-221607P.
PR 02-MAY-2001; 2001US-287724P.
XX
PA (COMP-) COMPUGEN INC.
XX
PI Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
XX
DR WPI; 2002-257383/30.
XX
PT New oligonucleotide libraries comprising oligonucleotides which
PT selectively hybridize to mRNAs transcribed from a transcription unit of
PT a genome, useful for detecting tissue-, pathology-, and
PT developmental-specific genes -
XX
PS Example 1; SEQ ID 7908; 47pp; English.
XX
CC The present invention describes oligonucleotide libraries for detecting
CC messenger RNAs that populate a (sub-)transcriptome, where the
CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC transcription units that populate a genome. The library comprises
CC several oligonucleotides, each capable of hybridising selectively to a
CC set of messenger RNAs transcribed from a given transcription unit of
CC the genome, which encodes one or more messenger RNA splice variants.
CC The oligonucleotide libraries are useful for detecting mRNAs from a
CC biological sample, in expression profiling studies, in qualitatively or
CC quantitatively characterising the corresponding transcriptome, and in
CC detecting RNA transcripts and splice variants of human or animal
CC transcriptomes. The libraries may also be used as specialised mini
CC libraries to detect transcripts of a sub-transcriptome under a
CC particular biological or pathological state, and so allowing the
CC detection of tissue- and pathology-specific genes such as those genes
CC only expressed in specific tissue under a specific pathological
CC condition; to detect developmental specific genes; and to detect RNA
CC transcripts and splice variants of a transcriptome of a patient suffering
CC from a particular disorder. ABN27253 to ABN59589 represent
CC oligonucleotide sequences from rats, humans and mice, which are used in
CC the exemplification of the present invention.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 60 BP; 14 A; 17 C; 14 G; 15 T; 0 other;

Query Match 74.0%; Score 14.8; DB 24; Length 60;

```
  Best Local Similarity   88.9%;  Pred. No. 1.4e+03;
  Matches   16;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       3 TTGCCTGGCTTCCTGGAG 20
           |||| |||| ||||||||
Db      41 TTGCATGGCATCCTGGAG 24


RESULT 6
AAC14135
ID   AAC14135 standard; cDNA; 87 BP.
XX
AC   AAC14135;
XX
DT   06-OCT-2000  (first entry)
XX
DE   Human secreted protein 5' EST, SEQ ID NO: 18210.
XX
KW   Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW   gene therapy; chromosome mapping; ss.
XX
OS   Homo sapiens.
XX
PN   EP1033401-A2.
XX
PD   06-SEP-2000.
XX
PF   21-FEB-2000; 2000EP-0200610.
XX
PR   26-FEB-1999;   99US-0122487.
XX
PA   (GEST ) GENSET.
XX
PI   Dumas Milne Edwards J,  Duclert A,  Giordano J;
XX
DR   WPI; 2000-500381/45.
XX
PT   New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT   obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT   diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX
PS   Claim 1; SEQ ID 18210; 71pp + CD-ROM; English.
XX
CC   The present sequence is one of a large number of 5' ESTs derived from
CC   mRNAs encoding secreted proteins. No ORF has yet been conclusively
CC   identified within the present sequence. The 5' ESTs were prepared from
CC   total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
CC   sequences usually correspond mainly to the 3' untranslated region (UTR)
CC   of the mRNA because they are often obtained from oligo-dT primed cDNA
CC   libraries. Such ESTs are not well suited for isolating cDNA sequences
CC   derived from the 5' ends of mRNAs and even in those cases where longer
CC   cDNA sequences have been obtained, the full 5' UTR is rarely included.
CC   5' ESTs are derived from mRNAs with intact 5' ends and can therefore be
CC   used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used
CC   in diagnostic, forensic, gene therapy and chromosome mapping procedures.
CC   They are used to obtain upstream regulatory sequences and to design
CC   expression and secretion vectors.
XX
SQ   Sequence 87 BP; 15 A; 20 C; 25 G; 27 T; 0 other;

  Query Match             74.0%;  Score 14.8;  DB 21;  Length 87;
  Best Local Similarity   88.9%;  Pred. No. 1.5e+03;
  Matches   16;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       3 TTGCCTGGCTTCCTGGAG 20
           | ||||||||||||| ||
Db      27 TGGCCTGGCTTCCTGCAG 44


RESULT 7
AAZ50174/c
ID   AAZ50174 standard; DNA; 26 BP.
XX
AC   AAZ50174;
XX
DT   04-MAY-2000  (first entry)
XX
DE   PCR primer DAE6 for raf gene amplification.
XX
KW   PCR primer; modulator; protein function; gene therapy; raf gene;
KW   tagged dominant negative element; TDNE; metabolic disorder;
KW   proliferative disorder; endocrine disorder; diabetes; obesity; porphyria;
KW   phenylketonuria; arrhythmia; Raf interaction; ss.
XX
OS   Saccharomyces cerevisiae.
XX
PN   WO200005417-A1.
XX
PD   03-FEB-2000.
XX
PF   23-JUL-1999;   99WO-US16749.
XX
PR   23-JUL-1998;   98US-0093855.
XX
PA   (SMAL-) SMALL MOLECULE THERAPEUTICS INC.
XX
PI   Menzel R,  Khazak V;
XX
DR   WPI; 2000-182729/16.
XX
PT   Identification modulatory elements of functional protein interaction
PT   for use in determining biological activity of the protein and for
PT   treating diseases  -
XX
PS   Example 1; Page 66; 104pp; English.
XX
CC   The patent discloses methods for determination of protein function and
CC   identifying modulators. This involves use of tagged dominant negative
CC   element (TDNE) that interferes with the interaction between a target and
CC   a partner protein comprising expressing TDNE in a microbial cell,
CC   measuring reporter gene expression and comparing the level to a
CC   expression level obtained in the absence of TDNE. The TDNEs are used in
CC   gene therapy and in screening for modulators. This is used for treating
CC   metabolic, proliferative, or endocrine disorders like diabetes, obesity,
CC   porphyria, phenylketonuria and arrhythmias. The present sequence is PCR
CC   primer DAE6 for amplification of full length raf gene and introduction
CC   of PstI and BamHI restriction sites. It is used in the isolation of
CC   interaction-based Raf blocking TDNE in the E. coli AraC system. The
CC   plasmids are used to verify the ability of engineered AraC fusion
CC   protein to produce a detectable change in the system.
XX
SQ   Sequence 26 BP; 8 A; 5 C; 9 G; 4 T; 0 other;

  Query Match             72.0%;  Score 14.4;  DB 21;  Length 26;
  Best Local Similarity   93.8%;  Pred. No. 2e+03;
  Matches   15;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       4 TGCCTGGCTTCCTGGA 19
           |||||| |||||||||
Db      25 TGCCTGTCTTCCTGGA 10


RESULT 8
AAL29599/c
ID   AAL29599 standard; DNA; 51 BP.
XX
AC   AAL29599;
XX
DT   24-JAN-2002  (first entry)
XX
DE   Human SNP oligonucleotide #2807.
XX
KW   Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
KW   neuroprotective; antimicrobial; gene therapy; vaccine; amylase; cancer;
KW   amyloid protein; angiopoietin; apoptosis related protein; cadherin;
KW   cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
```

KW complement related protein; cytochrome; kinesin; cytokine; interferon;
KW interleukin; G-protein coupled receptor; thioesterase; inflammation;
KW multifactorial disease; autoimmune disease; infection;
KW nervous system disease; ss.
XX
OS Homo sapiens.
XX
PN WO200147944-A2.
XX
PD 05-JUL-2001.
XX
PF 28-DEC-2000; 2000WO-US35498.
XX
PR 28-DEC-1999; 99US-0173419.
PR 27-DEC-2000; 2000US-0173419.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shimkets RA, Leach M;
XX
DR WPI; 2001-465210/50.
XX
PT Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
PT oncogenes and histones, useful for diagnosing and treating, e.g.
PT cancer, autoimmune diseases and infections -
XX
PS Claim 1; Page 2189; 4143pp; English.
XX
CC The present invention relates to oligonucleotides encoding polymorphic
CC variants of proteins related to amylases, amyloid proteins, angiopoietin,
CC apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,
CC histones, kinases, colony stimulating factors, complement related
CC proteins, cytochromes, kinesins, cytokines, interferons, interleukins,
CC G-protein coupled receptors and thioesterases. The present sequence is
CC one such oligonucleotide. The oligonucleotides and the peptides encoded
CC by them may be used in the prevention, diagnosis and treatment of
CC diseases associated with inappropriate expression of the proteins listed
CC above. Disorders that may be prevented, diagnosed and/or treated include
CC multifactorial diseases with a genetic component, such as autoimmune
CC diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,
CC systemic lupus erythromatosus and Grave's disease), inflammation, cancer
CC (e.g. cancers of the bladder, brain, breast, colon and kidney,
CC leukaemia), diseases of the nervous system and an infection of pathogenic
CC organisms.
XX
SQ Sequence 51 BP; 13 A; 16 C; 14 G; 8 T; 0 other;

Query Match 72.0%; Score 14.4; DB 22; Length 51;
Best Local Similarity 93.8%; Pred. No. 2.2e+03;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 5 GCCTGGCTTCCTGGAG 20
||||||||||| ||||
Db 32 GCCTGGCTTCCAGGAG 17

RESULT 9
ABN56037/c
ID ABN56037 standard; DNA; 65 BP.
XX
AC ABN56037;
XX
DT 15-JUL-2002 (first entry)
XX
DE Mouse spliced transcript detection oligonucleotide SEQ ID NO:28785.
XX
KW Human; mouse; rat; splice transcript; detection; RNA transcript;
KW splice variant; transcriptome; oligonucleotide library; ss.
XX
OS Mus musculus.
XX
PN WO200210449-A2.
XX
PD 07-FEB-2002.
XX
PF 20-JUL-2001; 2001WO-IB01903.
XX
PR 28-JUL-2000; 2000US-221607P.
PR 02-MAY-2001; 2001US-287724P.
XX
PA (COMP-) COMPUGEN INC.
XX
PI Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
XX
DR WPI; 2002-257383/30.
XX
PT New oligonucleotide libraries comprising oligonucleotides which
PT selectively hybridize to mRNAs transcribed from a transcription unit of
PT a genome, useful for detecting tissue-, pathology-, and
PT developmental-specific genes -
XX
PS Example 1; SEQ ID 28785; 47pp; English.
XX
CC The present invention describes oligonucleotide libraries for detecting
CC messenger RNAs that populate a (sub-)transcriptome, where the
CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC transcription units that populate a genome. The library comprises
CC several oligonucleotides, each capable of hybridising selectively to a
CC set of messenger RNAs transcribed from a given transcription unit of
CC the genome, which encodes one or more messenger RNA splice variants.
CC The oligonucleotide libraries are useful for detecting mRNAs from a
CC biological sample, in expression profiling studies, in qualitatively or
CC quantitatively characterising the corresponding transcriptome, and in
CC detecting RNA transcripts and splice variants of human or animal
CC transcriptomes. The libraries may also be used as specialised mini
CC libraries to detect transcripts of a sub-transcriptome under a
CC particular biological or pathological state, and so allowing the
CC detection of tissue- and pathology-specific genes such as those genes
CC only expressed in specific tissue under a specific pathological
CC condition; to detect developmental specific genes; and to detect RNA
CC transcripts and splice variants of a transcriptome of a patient suffering
CC from a particular disorder. ABN27253 to ABN59589 represent
CC oligonucleotide sequences from rats, humans and mice, which are used in
CC the exemplification of the present invention.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 65 BP; 16 A; 17 C; 16 G; 16 T; 0 other;

Query Match 72.0%; Score 14.4; DB 24; Length 65;
Best Local Similarity 93.8%; Pred. No. 2.2e+03;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CATTGCCTGGCTTCCT 16
|| ||||||||||||||
Db 56 CAGTGCCTGGCTTCCT 41

RESULT 10
AAT32566/c
ID AAT32566 standard; DNA; 36 BP.
XX
AC AAT32566;
XX
DT 14-JAN-1997 (first entry)
XX
DE E.coli polyphosphate kinase gene amplification primer.
XX
KW Polyphosphate kinase; PPK; acetate kinase; phosphotransacetylase;
KW plant vacuole; phosphate availability; mobilisation; crop yield;
KW increase; transgenic plant; potato; polymerase chain reaction;
KW PCR primer; early flowering; ss.
XX
OS Synthetic.
XX

PN DE4444460-A1.
XX
PD 30-MAY-1996.
XX
PF 29-NOV-1994; 94DE-4444460.
XX
PR 29-NOV-1994; 94DE-4444460.
XX
PA (GENB-) INST GENBIOLOGISCHE FORSCHUNG.
XX
PI Kossmann J, Springer F;
XX
DR WPI; 1996-260704/27.
XX
PT Increasing biomass prodn. in plants by creating phosphate pool
PT outside the vacuole for easy mobilisation - useful also for
PT simultaneous modification of flowering time
XX
PS Example 1; Page 9; 15pp; German.
XX
CC A PCR primer of the present sequence was used for PCR amplification
CC of the ppk gene of E.coli, coding for the polyphosphate kinase (PPK)
CC enzyme. The amplified fragment was incorporated into a plasmid for
CC expression of PPK in plant cells under transcriptional control of
CC the cauliflower mosaic virus 35S promoter and the terminator from
CC Ti plasmid T-DNA gene3. The vector was used for generating
CC transgenic potato plants which expressed PPK; these plants were
CC shown to flower earlier than wild-type plants and to produce a
CC greater yield of potato tubers. Also claimed are transgenic plants
CC which have the acetate kinase or phosphotransacetylase gene, in stead
CC of the ppk gene, integrated into their genomes; all these plants are
CC able to produce acetyl phosphate or polyphosphate outside the
CC vacuole so that there is a pool of easily mobilised phosphate.
CC Biomass production in the plants increases as a result of increased
CC phosphate availability.
XX
SQ Sequence 36 BP; 12 A; 8 C; 13 G; 3 T; 0 other;

Query Match 71.0%; Score 14.2; DB 17; Length 36;
Best Local Similarity 84.2%; Pred. No. 2.6e+03;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CATTGCCTGGCTTCCTGGA 19
||| ||| || ||||||||
Db 24 CATGGCCCGGGTTCCTGGA 6

RESULT 11
ABK65538
ID ABK65538 standard; DNA; 21 BP.
XX
AC ABK65538;
XX
DT 02-JUL-2002 (first entry)
XX
DE Human single nucleotide polymorphism #158.
XX
KW Human; single nucleotide polymorphism; SNP; sickle cell anaemia;
KW agammaglobulinaemia; diabetes insipidus; Lesch-Nyhan syndrome;
KW muscular dystrophy; Wiskott-Aldrich syndrome; Fabry's disease;
KW familial hypercholesterolaemia; polycystic kidney disease; cancer;
KW hereditary spherocytosis; Von Willebrand's disease; tuberous sclerosis;
KW hereditary haemorrhagic telangiectasia; familial colonic polyposis;
KW Ehlers-Danlos syndrome; osteogenesis imperfecta; autoimmune disease;
KW acute intermittent porphyria; inflammation; nervous system disorder;
KW infection; rheumatoid arthritis; multiple sclerosis; diabetes;
KW systemic lupus erythematosus; Graves disease; longevity; obesity;
KW baldness; fertility; forensic; paternity testing; ss.
XX
OS Homo sapiens.
XX
PN US2002037508-A1.
XX
PD 28-MAR-2002.
XX
PF 18-JAN-2001; 2001US-0765081.
XX
PR 19-JAN-2000; 2000US-176861P.
XX
PA (CARG/) CARGILL M.
PA (IREL/) IRELAND J S.
PA (LAND/) LANDER E S.
XX
PI Cargill M, Ireland JS, Lander ES;
XX
DR WPI; 2002-315108/35.
XX
PT Nucleic acid comprising single nucleotide polymorphisms, useful in
PT forensics, paternity testing and diagnosis of disease -
XX
PS Claim 1; Page 54; 96pp; English.
XX
CC The invention relates to a nucleic acid comprising single nucleotide
CC polymorphisms (SNPs) associated with diseases. The nucleic acids
CC comprising the SNPs and probes and primers for detecting them may be used
CC in assays for the diagnosis of diseases associated with SNPs (such as
CC sickle cell anaemia, agammaglobulinaemia, diabetes insipidus, Lesch-Nyhan
CC syndrome, muscular dystrophy, Wiskott-Aldrich syndrome, Fabry's disease,
CC familial hypercholesterolaemia, polycystic kidney disease, hereditary
CC spherocytosis, Von Willebrand's disease, tuberous sclerosis, hereditary
CC haemorrhagic telangiectasia, familial colonic polyposis, Ehlers-Danlos
CC syndrome, osteogenesis imperfecta, and acute intermittent porphyria,
CC symptoms of, or susceptibility to, multifactorial diseases of which a
CC component is or may be genetic, such as autoimmune diseases,
CC inflammation, cancer, diseases of the nervous system, and infection by
CC pathogenic microorganisms, autoimmune diseases including rheumatoid
CC arthritis, multiple sclerosis, diabetes (insulin-dependent and
CC non-independent), systemic lupus erythematosus and Graves disease,
CC cancers including cancers of the bladder, brain, breast, colon,
CC oesophagus, kidney, leukaemia, liver, lung, oral cavity, ovary, pancreas,
CC prostate, skin, stomach and uterus, longevity, appearance (e.g.,
CC baldness, obesity), strength, speed, endurance, fertility, and
CC susceptibility or receptivity to particular drugs or therapeutic
CC treatments), in forensics and in paternity testing. ABK65381-ABK65841
CC represent human single nucleotide polymorphisms of the invention.
XX
SQ Sequence 21 BP; 5 A; 5 C; 6 G; 4 T; 1 other;

Query Match 70.0%; Score 14; DB 24; Length 21;
Best Local Similarity 87.5%; Pred. No. 3.1e+03;
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 4 TGCCTGGCTTCCTGGA 19
||||||||| |||||
Db 2 TGCCTGGCTYACTGGA 17

RESULT 12
AAH26781
ID AAH26781 standard; DNA; 24 BP.
XX
AC AAH26781;
XX
DT 26-NOV-2001 (first entry)
XX
DE Mouse T cell receptor V-alpha chain PCR primer V-alpha 5T.
XX
KW T cell receptor; mouse; antigen; atherosclerosis; vaccine;
KW diagnosis; PCR primer; ss.
XX
OS Mus sp.
XX
PN WO200168119-A1.
XX
PD 20-SEP-2001.
XX

PF 15-MAR-2001; 2001WO-SE00570.
XX
PR 15-MAR-2000; 2000SE-0000855.
XX
PA (KARO-) KAROLINSKA INNOVATIONS AB.
XX
PI Hansson GK, Stemme S, Nicoletti A, Wuttge D;
XX
DR WPI; 2001-589990/66.
XX
PT Antigenic composition useful as an atherosclerosis vaccine comprises a
PT peptide derived from apoB100 conjugated with an aldehyde, and is
PT capable of eliciting an immune response against non-native, oxidized
PT low density lipoprotein -
XX
PS Disclosure; Page 23; 49pp; English.
XX
CC The present sequence is that of primer V-alpha 5T, which is 1 of
CC a set of 19 V-alpha primers (see AAH26763-81) used with a C-alpha
CC primer (see AAH26782) for the PCR amplification of mouse T cell
CC receptor (TcR) V-alpha chain cDNA. The TcR from hybridoma 96-6,
CC which was reactive to oxidised low density lipoprotein (oxLDL),
CC carried a V-beta-6 chain (see AAB82887) with a cluster of charged
CC and polar amino acids in complementarity determining region 3.
CC The hybridoma recognised purified apoB100 protein conjugated to
CC malondialdehyde but not the native, unconjugated apoB100 protein.
CC The invention provides antigenic compositions comprising a peptide
CC derived from apoB100, conjugated with an aldehyde, and capable of
CC eliciting an immune response against non-native, oxLDL in a subject
CC by interacting with TcRs, especially with TcR alpha-10 and beta-6
CC chains. The compositions are used as vaccines against
CC atherosclerosis. Methods of producing an atherosclerosis vaccine
CC and in diagnosing atherosclerosis are also claimed.
XX
SQ Sequence 24 BP; 5 A; 7 C; 7 G; 5 T; 0 other;

Query Match 69.0%; Score 13.8; DB 22; Length 24;
Best Local Similarity 88.2%; Pred. No. 3.8e+03;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 TGCCTGGCTTCCTGGAG 20
|| |||||||||| ||
Db 5 TGACTGGCTTCCTGAAG 21

RESULT 13
AAA35734
ID AAA35734 standard; DNA; 36 BP.
XX
AC AAA35734;
XX
DT 26-JUL-2000 (first entry)
XX
DE Permutein linker encoding nucleotide sequence FGS159.
XX
KW Biologically-activated circularly-permuted protein; permutein; linker;
KW permutein library generation; therapeutic property; antigen;
KW immunotherapy; improve bio-distribution; half life; ss.
XX
OS Synthetic.
XX
PN WO200018905-A1.
XX
PD 06-APR-2000.
XX
PF 24-SEP-1999; 99WO-US20891.
XX
PR 25-SEP-1998; 98US-0101908.
XX
PA (SEAR) SEARLE & CO G D.
XX
PI Lee SC;
XX
DR WPI; 2000-293145/25.
XX
PT Preparation of biologically-activated circularly-permuted proteins by
PT scanning permutagenesis for generating libraries of permuteins with
PT improved therapeutic properties -
XX
PS Claim 11; Page 41; 100pp; English.
XX
CC The preparation of biologically-activated circularly-permuted proteins
CC (permuteins) comprises the use of a method comprising making a series of
CC circularly permuted genes. The circularly permuted genes are inserted
CC into a display vector, where they are expressed so that the proteins
CC they encode are presented on the surface of the display vector. A library
CC of display vectors presenting the expressed circularly permuted proteins
CC is generated. A target protein that can bind a biologically active
CC circularly permuted protein can be used to affinity select the
CC presenting display vectors. The selected display vectors can be isolated
CC and analysed to identify the presented circularly-permuted protein. The
CC permuteins conform to the structure of a parent protein consisting of a
CC segment derived from the carboxy portion of the parent protein, a
CC segment derived from the amino terminus of the parent protein, and a
CC linker or chemical bond linking the amino and carboxy terminal derived
CC portions. Nucleotide sequences AAA35576-A35943 encode linkers used to
CC create the permuteins of the invention. The method is used to generate
CC libraries of permuteins with improved therapeutic properties compared to
CC their parent molecules. Permuteins with little or no activity may be used
CC as antigens for producing antibodies which are used in immunology or
CC immunotherapy as probes or intermediates used to construct other useful
CC permuteins. Permuteins have improved biological and therapeutic
CC properties compared to their two individual components due to alterations
CC in bio-distribution or half-life.
XX
SQ Sequence 36 BP; 5 A; 11 C; 11 G; 9 T; 0 other;

Query Match 69.0%; Score 13.8; DB 21; Length 36;
Best Local Similarity 88.2%; Pred. No. 4e+03;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 TGCCTGGCTTCCTGGAG 20
|| || ||||||||||
Db 11 TGTCTAGCTTCCTGGAG 27

RESULT 14
AAZ96894/c
ID AAZ96894 standard; DNA; 59 BP.
XX
AC AAZ96894;
XX
DT 14-APR-2000 (first entry)
XX
DE S. cerevisiae gene deletion cassette constructing primer YML127w-S1.
XX
KW Antimycotic; mycosis; immunodepression; AIDS; diabetes; fungicide;
KW mycete; gene deletion; PCR primer; ss.
XX
OS Saccharomyces cerevisiae.
XX
PN WO9955907-A2.
XX
PD 04-NOV-1999.
XX
PF 22-APR-1999; 99WO-EP02722.
XX
PR 24-APR-1998; 98EP-0401007.
PR 11-SEP-1998; 98EP-0402254.
XX
PA (HMRI) HOECHST MARION ROUSSEL.
XX
PI Diu-Hercend A, Entian K, Koetter P;
XX
DR WPI; 2000-105527/09.
XX

PT Identifying antimycotic substances useful for drug preparation and
PT treatment of mycosis -
XX
PS Examples; Page 65; 86pp; English.
XX
CC The invention provides a method of screening for antimycotic substances
CC using essential genes from mycetes or a functionally similar mycete
CC gene or the corresponding encoded protein as target. The essential gene
CC useful for screening antimycotic substances is selected from the
CC following genes: YML114c, YLR186w, YLR215c, YLR222c, YLR243w, YLR272c,
CC YLR275w, YLR276c, YLR317w, YLR359w, YLR373c, YLR424w, YLR437c, YLR440c,
CC YML023c, YML049c, YML077w, YML093w, YML127w, YMR032w, YMR093w, YMR131c,
CC YMR185w, YMR212c, YMR213w, YMR218c, YMR281w, YMR288w, YMR290c, YMR211w,
CC YMR049c, YMR134w, YDR196c, YDR299w, YDR365c, YDR396w, YDR407c, YDR416w,
CC YDR449c, YDR472w, YDR499w, YDR141c, YDR324c, YDR325w, YDR398w, YDR246w,
CC YDR236c, YDR361c, YDR367w, YDR339c, YDR413c, YDR429c, YDR468c, YDR489w,
CC YDR527w, YDR288w, YDR201w, YDR434w, YDR181c, YDR531w, YPL126w, YPL093w,
CC YPL063w, YPL024w, YPL020c, YPL012w, YPL007c, YPL233w, YPL146c, YIL091c,
CC YIL083c, YIL019w, YIL109c, YIL104c, YFL024c, YFR003c, YFR027w, YFR042w,
CC YIR010w, YIR015w, YPR048w, YPR072w, YPR082c, YPR085c, YPR105c, YPR112c,
CC YPR137w, YPR143w, YPR144c and YPR169w. The method is useful for
CC identifying substances for the preparation of drugs for the treatment of
CC mycosis or prevention in immunodepression states. Drugs containing
CC antimycotic substances are useful for the treatment of mycotic
CC infections which occur during diseases like AIDS or diabetes. Substances
CC which may be used for the fabrication of fungicides, especially of
CC fungicides which are harmless for humans and animals and antimycotic
CC substances which selectively inhibit the growth of specific mycete
CC species only, can also be identified by this method. Sequences
CC AAZ96811-Z96990 represent PCR primers used in construction of S.
CC cerevisiae deletion cassettes.
XX
SQ Sequence 59 BP; 16 A; 20 C; 14 G; 9 T; 0 other;

Query Match 69.0%; Score 13.8; DB 21; Length 59;
Best Local Similarity 88.2%; Pred. No. 4.2e+03;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy       4 TGCCTGGCTTCCTGGAG 20
           ||||| |||| ||||||
Db      30 TGCCTCGCTTACTGGAG 14
```

RESULT 15
AAC60378
ID AAC60378 standard; DNA; 60 BP.
XX
AC AAC60378;
XX
DT 16-FEB-2001 (first entry)
XX
DE DNA encoding peptide leader sequence.
XX
KW Recombinant protein; E1; ss.
XX
OS Synthetic.
XX
PN WO200063403-A2.
XX
PD 26-OCT-2000.
XX
PF 17-APR-2000; 2000WO-NL00247.
XX
PR 15-APR-1999; 99EP-0201176.
PR 21-DEC-1999; 99EP-0204434.
XX
PA (INTR-) INTROGENE BV.
XX
PI Hateboer G, Verhulst KC, Schouten GJ, Uytdehaag AGCM, Bout A;
XX
DR WPI; 2000-665247/64.
XX
PT Producing recombinant proteins, e.g. viral proteins for use in
PT vaccines, in a human cell line which encodes at least one E1 protein
PT but does encode structural adenoviral proteins -
XX
PS Example 3; Page 32; 127pp; English.
XX
CC The present invention relates to a method of producing recombinant
CC proteins in a eukaryotic cell, preferably a human cell line. The method
CC involves using a sequence encoding at least one E1 protein of an
CC adenovirus or its functional homologue, fragment or derivative.
CC The cell does not encode a structural adenoviral protein from its
CC genome or from a sequence integrated into the genome. The recombinant
CC mammalian cells are useful for producing variable domains of an
CC immunoglobulin which have post-translational modifications different
CC than that of their natural counterparts. The cell is capable of
CC producing 2 to 200 fold more recombinant protein and/or proteinaceous
CC substances than conventional mammalian lines
XX
SQ Sequence 60 BP; 8 A; 17 C; 14 G; 21 T; 0 other;

Query Match 69.0%; Score 13.8; DB 21; Length 60;
Best Local Similarity 88.2%; Pred. No. 4.2e+03;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy       1 CATTGCCTGGCTTCCTG 17
           |||  ||||||||||||
Db       5 CATGCCCTGGCTTCCTG 21
```

Search completed: November 23, 2002, 06:29:30
Job time : 102.6 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 05:54:41 ; Search time 17.25 Seconds
(without alignments)
439.108 Million cell updates/sec

Title: US-09-296-264-19
Perfect score: 20
Sequence: 1 cattgcctggcttcctggag 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 335578 seqs, 189365133 residues

Total number of hits satisfying chosen parameters: 195614

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	14	70.0	21	10	US-09-765-081-158	Sequence 158, App
2	13.8	69.0	60	10	US-09-940-386-2	Sequence 2, Appli
c 3	13.2	66.0	24	9	US-10-036-492-24	Sequence 24, Appl
4	13.2	66.0	86	10	US-09-215-652-4	Sequence 4, Appli
5	12.8	64.0	30	10	US-09-941-094A-3	Sequence 3, Appli
c 6	12.8	64.0	36	10	US-09-855-722-31	Sequence 31, Appl
c 7	12.8	64.0	37	10	US-09-995-593A-23	Sequence 23, Appl
8	12.8	64.0	98	10	US-09-864-761-26125	Sequence 26125, A
c 9	12.8	64.0	99	10	US-09-864-761-21642	Sequence 21642, A
10	12.6	63.0	32	9	US-09-840-243B-4	Sequence 4, Appli
c 11	12.6	63.0	60	10	US-09-735-705-370	Sequence 370, App
c 12	12.6	63.0	60	10	US-09-850-716A-370	Sequence 370, App
c 13	12.6	63.0	60	10	US-09-897-778-370	Sequence 370, App
c 14	12.6	63.0	99	10	US-09-864-761-33144	Sequence 33144, A
15	12.6	63.0	100	10	US-09-294-093B-1610	Sequence 1610, Ap
16	12.4	62.0	24	9	US-09-943-388-25	Sequence 25, Appl
c 17	12.4	62.0	25	10	US-09-828-313-114	Sequence 114, App
c 18	12.4	62.0	98	10	US-09-864-761-22944	Sequence 22944, A
19	12.2	61.0	18	10	US-09-067-638B-69	Sequence 69, Appl
c 20	12.2	61.0	31	9	US-09-773-599-1	Sequence 1, Appli
21	12.2	61.0	65	9	US-10-004-201-31	Sequence 31, Appl
22	12.2	61.0	80	10	US-09-864-761-26919	Sequence 26919, A
c 23	12.2	61.0	87	10	US-09-864-761-20910	Sequence 20910, A
24	12	60.0	31	10	US-09-801-274-122	Sequence 122, App
c 25	12	60.0	36	10	US-09-986-666-6	Sequence 6, Appli
c 26	12	60.0	36	10	US-09-986-667-6	Sequence 6, Appli
c 27	12	60.0	39	9	US-10-157-855-34	Sequence 34, Appl
c 28	12	60.0	84	10	US-09-921-397-49	Sequence 49, Appl
c 29	12	60.0	91	10	US-09-864-761-18106	Sequence 18106, A
c 30	12	60.0	95	10	US-09-864-761-29634	Sequence 29634, A
31	11.8	59.0	17	10	US-09-866-108-9548	Sequence 9548, Ap
32	11.8	59.0	17	10	US-09-866-108-9549	Sequence 9549, Ap
33	11.8	59.0	17	10	US-09-866-108-9550	Sequence 9550, Ap
c 34	11.8	59.0	19	10	US-09-814-986-39	Sequence 39, Appl
35	11.8	59.0	21	9	US-09-966-546-49	Sequence 49, Appl
36	11.8	59.0	25	10	US-09-866-108-14440	Sequence 14440, A
37	11.8	59.0	25	10	US-09-866-108-14441	Sequence 14441, A
38	11.8	59.0	25	10	US-09-866-108-14442	Sequence 14442, A
39	11.8	59.0	25	10	US-09-866-108-14443	Sequence 14443, A
40	11.8	59.0	25	10	US-09-866-108-14444	Sequence 14444, A
41	11.8	59.0	25	10	US-09-866-108-14445	Sequence 14445, A
42	11.8	59.0	25	10	US-09-866-108-14446	Sequence 14446, A
43	11.8	59.0	25	10	US-09-866-108-14447	Sequence 14447, A
44	11.8	59.0	25	10	US-09-866-108-14448	Sequence 14448, A
45	11.8	59.0	25	10	US-09-866-108-14449	Sequence 14449, A

ALIGNMENTS

RESULT 1
US-09-765-081-158
; Sequence 158, Application US/09765081
; Patent No. US20020037508A1
; GENERAL INFORMATION:
; APPLICANT: Cargill, Michele
; APPLICANT: Ireland, James S.
; APPLICANT: Lander, Eric S.
; TITLE OF INVENTION: HUMAN SINGLE NUCLEOTIDE POLYMORPHISMS
; FILE REFERENCE: 2825.2008-001
; CURRENT APPLICATION NUMBER: US/09/765,081
; CURRENT FILING DATE: 2001-01-18
; PRIOR APPLICATION NUMBER: US 60/176,861
; PRIOR FILING DATE: 2000-01-19
; NUMBER OF SEQ ID NOS: 461
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 158
; LENGTH: 21
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-765-081-158

Query Match 70.0%; Score 14; DB 10; Length 21;
Best Local Similarity 87.5%; Pred. No. 4.2e+02;
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```
Qy       4 TGCCTGGCTTCCTGGA 19
           |||||||||: |||||
Db       2 TGCCTGGCTYACTGGA 17
```

RESULT 2
US-09-940-386-2
; Sequence 2, Application US/09940386
; Patent No. US20020115065A1
; GENERAL INFORMATION:
; APPLICANT: Logtenberg, Ton
; APPLICANT: Cilenti, Lucia
; APPLICANT: Bloem, Andries C.
; APPLICANT: Zwijsen, Renate M.L.
; TITLE OF INVENTION: Differentially expressed epitopes and uses thereof
; FILE REFERENCE: 2183-4514.1US

```
;  CURRENT APPLICATION NUMBER: US/09/940,386
;  CURRENT FILING DATE:  2001-08-27
;  PRIOR APPLICATION NUMBER: EP 00202991.6
;  PRIOR FILING DATE: 2000-08-28
;  PRIOR APPLICATION NUMBER: US 60/228.429
;  PRIOR FILING DATE: 2000-08-28
;  NUMBER OF SEQ ID NOS: 11
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
;   LENGTH: 60
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: part of plasmid pNUT-c gamma-
;   OTHER INFORMATION: UBS-54 encoding leader peptide sequence
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(60)
;   OTHER INFORMATION:
US-09-940-386-2

  Query Match             69.0%;  Score 13.8;  DB 10;  Length 60;
  Best Local Similarity   88.2%;  Pred. No. 5.9e+02;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 CATTGCCTGGCTTCCTG 17
           |||  ||||||||||||
Db       5 CATGCCCTGGCTTCCTG 21


RESULT 3
US-10-036-492-24/c
; Sequence 24, Application US/10036492
; Patent No. US20020164757A1
; GENERAL INFORMATION:
;  APPLICANT: HEMERLY, ADRIANA
;  APPLICANT:  FERREIRA, PAULO
;  APPLICANT:  ROMBAUTS, STEPHANE
;  TITLE OF INVENTION: PLANT DNA REPLICATION MODULATING PROTEINS
;  FILE REFERENCE: 217943US0XCONT
;  CURRENT APPLICATION NUMBER: US/10/036,492
;  CURRENT FILING DATE:  2002-01-07
;  PRIOR APPLICATION NUMBER: EP 99202214.5
;  PRIOR FILING DATE: 1999-07-05
;  NUMBER OF SEQ ID NOS: 35
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 24
;   LENGTH: 24
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Synthetic DNA
US-10-036-492-24

  Query Match             66.0%;  Score 13.2;  DB 9;  Length 24;
  Best Local Similarity   83.3%;  Pred. No. 1e+03;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       3 TTGCCTGGCTTCCTGGAG 20
           ||||||  |||||| |||
Db      21 TTGCCTTTCTTCCTCGAG 4


RESULT 4
US-09-215-652-4
; Sequence 4, Application US/09215652
; Patent No. US20020045165A1
; GENERAL INFORMATION:
;  APPLICANT: Patricia Billing-Medel
;  APPLICANT: Maurice Cohen
;  APPLICANT: Tracey L. Colpitts
;  APPLICANT: Paula N. Friedman
;  APPLICANT: Julian Gordon
;  APPLICANT: Edward N. Granados
;  APPLICANT: Steven C. Hodges
;  APPLICANT: Michael R. Klass
;  APPLICANT: Jon D. Kratochvil
;  APPLICANT: Lisa Roberts-Rapp
;  APPLICANT: John C. Russell
;  APPLICANT: Stephen D. Stroupe
;  TITLE OF INVENTION: Reagents and Methods Useful for Detecting Diseases of the
;  TITLE OF INVENTION: Breast
;  FILE REFERENCE: 6192.US.P1
;  CURRENT APPLICATION NUMBER: US/09/215,652
;  CURRENT FILING DATE: 1998-12-16
;  EARLIER APPLICATION NUMBER: US 08/998,496
;  EARLIER FILING DATE: 1997-12-26
;  NUMBER OF SEQ ID NOS: 48
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 4
;   LENGTH: 86
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: base_polymorphism
;   LOCATION: 56
;   OTHER INFORMATION: /note = " n' represents an a or g or t or c polymorphism at
;   OTHER INFORMATION: this position
US-09-215-652-4

  Query Match             66.0%;  Score 13.2;  DB 10;  Length 86;
  Best Local Similarity   83.3%;  Pred. No. 1.2e+03;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       3 TTGCCTGGCTTCCTGGAG 20
           | || ||||| |||||||
Db      18 TAGCTTGGCTGCCTGGAG 35


RESULT 5
US-09-941-094A-3
; Sequence 3, Application US/09941094A
; Patent No. US20020065226A1
; GENERAL INFORMATION:
;  APPLICANT: Siler-Khodr, Theresa M.
;  TITLE OF INVENTION: No. US20020065226A1-Mammalian GnRH Analogs and Uses Thereof in
;  TITLE OF INVENTION:  Pregnancy
;  FILE REFERENCE: P7345.2(CIP)
;  CURRENT APPLICATION NUMBER: US/09/941,094A
;  CURRENT FILING DATE:  2001-08-28
;  PRIOR APPLICATION NUMBER: US 09/419,161
;  PRIOR FILING DATE: 1999-10-15
;  NUMBER OF SEQ ID NOS: 4
; SEQ ID NO 3
;   LENGTH: 30
;   TYPE: DNA
;   ORGANISM: Salmo salar
US-09-941-094A-3

  Query Match             64.0%;  Score 12.8;  DB 10;  Length 30;
  Best Local Similarity   87.5%;  Pred. No. 1.6e+03;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       4 TGCCTGGCTTCCTGGA 19
           || |||||| ||||||
Db      15 TGGCTGGCTGCCTGGA 30


RESULT 6
US-09-855-722-31/c
; Sequence 31, Application US/09855722
; Patent No. US20020049306A1
; GENERAL INFORMATION:
;  APPLICANT: Sakano, Seiji
;  APPLICANT:  Itoh, Akira
;  TITLE OF INVENTION: DIFFERENTIATION-SUPPRESSIVE POLYPEPTIDE
;  FILE REFERENCE: KP-8576
```

```
;  CURRENT APPLICATION NUMBER: US/09/855,722
;  CURRENT FILING DATE:  2001-05-16
;  PRIOR APPLICATION NUMBER: 09/214,278
;  PRIOR FILING DATE: 1999-01-26
;  NUMBER OF SEQ ID NOS: 32
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 31
;   LENGTH: 36
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: synthetic DNA
US-09-855-722-31

  Query Match             64.0%;  Score 12.8;  DB 10;  Length 36;
  Best Local Similarity   87.5%;  Pred. No. 1.7e+03;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       5 GCCTGGCTTCCTGGAG 20
           |||||||||||  | ||
Db      33 GCCTGGCTTCCAGCAG 18


RESULT 7
US-09-995-593A-23/c
; Sequence 23, Application US/09995593A
; Patent No. US20020128197A1
; GENERAL INFORMATION:
;  APPLICANT: SAKANO, SEIJI
;  APPLICANT:  ITOH, AKIRA
;  TITLE OF INVENTION: DIFFERENTIATION-SUPPRESSIVE POLYPEPTIDE
;  FILE REFERENCE: KP8447DIV
;  CURRENT APPLICATION NUMBER: US/09/995,593A
;  CURRENT FILING DATE:  2002-02-28
;  PRIOR APPLICATION NUMBER: 09/068,740
;  PRIOR FILING DATE: 1998-06-18
;  PRIOR APPLICATION NUMBER: JP 7-299611
;  PRIOR FILING DATE: 1995-11-17
;  PRIOR APPLICATION NUMBER: JP 7-311811
;  PRIOR FILING DATE: 1995-11-30
;  PRIOR APPLICATION NUMBER: PCT/JP96/03356
;  PRIOR FILING DATE: 1996-11-15
;  NUMBER OF SEQ ID NOS: 48
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 23
;   LENGTH: 37
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Primer
US-09-995-593A-23

  Query Match             64.0%;  Score 12.8;  DB 10;  Length 37;
  Best Local Similarity   87.5%;  Pred. No. 1.7e+03;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       5 GCCTGGCTTCCTGGAG 20
           |||||||||||  | ||
Db      34 GCCTGGCTTCCAGCAG 19


RESULT 8
US-09-864-761-26125
; Sequence 26125, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 26125
;   LENGTH: 98
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AL022318.2
;   OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL =   2
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.5
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.9
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.2
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3.4
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.7
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.9
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.2
;   OTHER INFORMATION: NT HIT: AF311912.1, EVALUE 3.20e+00
;   OTHER INFORMATION: EST_HUMAN HIT: BE179517.1, EVALUE 3.00e-10
US-09-864-761-26125

  Query Match             64.0%;  Score 12.8;  DB 10;  Length 98;
  Best Local Similarity   87.5%;  Pred. No. 1.9e+03;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       3 TTGCCTGGCTTCCTGG 18
           |||||||||  |||||
Db      33 TTGCCTGGCCACCTGG 48


RESULT 9
US-09-864-761-21642/c
; Sequence 21642, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
```

```
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 21642
;   LENGTH: 99
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AL031177.1
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.4
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.7
;   OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 2.5
;   OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.1
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.3
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.7
;   OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 2.3
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.3
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.1
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.3
;   OTHER INFORMATION: SWISSPROT HIT: Q14031, EVALUE 6.00e-14
;   OTHER INFORMATION: EST_HUMAN HIT: AA926629.1, EVALUE 3.10e-01
;   OTHER INFORMATION: NT HIT: U04845.1, EVALUE 4.00e-49
US-09-864-761-21642

  Query Match             64.0%;  Score 12.8;  DB 10;  Length 99;
  Best Local Similarity   87.5%;  Pred. No. 1.9e+03;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       5 GCCTGGCTTCCTGGAG 20
            |||||||||| ||| |
Db      86 GCCTGGCTTCATGGGG 71


RESULT 10
US-09-840-243B-4
; Sequence 4, Application US/09840243B
; Patent No. US20020156258A1
; GENERAL INFORMATION:
;  APPLICANT: MASTERNAK, Krzysztof
;  APPLICANT:  REITH, Walter
;  APPLICANT:  MACH, Bernard
;  TITLE OF INVENTION: New Transcription Factor of MHC Class II Genes, Substances
;  TITLE OF INVENTION:  Capable of Inhibiting This New Transcription Factor and
;  TITLE OF INVENTION:  Medical Uses of These Substances
;  FILE REFERENCE: 010830-117
;  CURRENT APPLICATION NUMBER: US/09/840,243B
;  CURRENT FILING DATE:  2001-04-24
;  PRIOR APPLICATION NUMBER: EP 98120085.0
;  PRIOR FILING DATE: 1998-10-24
;  NUMBER OF SEQ ID NOS: 22
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
;   LENGTH: 32
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:primer
US-09-840-243B-4

  Query Match             63.0%;  Score 12.6;  DB 9;  Length 32;
  Best Local Similarity   78.9%;  Pred. No. 2e+03;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 ATTGCCTGGCTTCCTGGAG 20
            | ||||||| |||| | ||
Db      13 AGTGCCTGGGTTCCAGCAG 31


RESULT 11
US-09-735-705-370/c
; Sequence 370, Application US/09735705
; Patent No. US20020052329A1
; GENERAL INFORMATION:
;  APPLICANT:  Wang, Tongtong
;  APPLICANT:  Fan, Liqun
;  APPLICANT:  Kalos, Michael D.
;  APPLICANT:  Bangur, Chaitanya S.
;  APPLICANT:  Hosken, Nancy
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Li, Samuel X.
;  APPLICANT:  Wang, Aijun
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Fanger, Neil
;  TITLE OF INVENTION:  COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE:  210121.455C14
;  CURRENT APPLICATION NUMBER: US/09/735,705
;  CURRENT FILING DATE:  2000-12-12
;  NUMBER OF SEQ ID NOS:  419
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 370
;   LENGTH: 60
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-735-705-370

  Query Match             63.0%;  Score 12.6;  DB 10;  Length 60;
  Best Local Similarity   78.9%;  Pred. No. 2.2e+03;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 CATTGCCTGGCTTCCTGGA 19
```

```
          ||| || || || ||||||
Db      60 CATAGCATGACTCCCTGGA 42


RESULT 12
US-09-850-716A-370/c
; Sequence 370, Application US/09850716A
; Patent No. US20020115139A1
; GENERAL INFORMATION:
;  APPLICANT: Kalos, Michael D.
;  APPLICANT:  McNeill, Patricia D.
;  APPLICANT:  Retter, Marc W.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE: 210121.455C15
;  CURRENT APPLICATION NUMBER: US/09/850,716A
;  CURRENT FILING DATE:  2001-05-07
;  NUMBER OF SEQ ID NOS: 440
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 370
;   LENGTH: 60
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-850-716A-370

  Query Match             63.0%;  Score 12.6;  DB 10;  Length 60;
  Best Local Similarity   78.9%;  Pred. No. 2.2e+03;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 CATTGCCTGGCTTCCTGGA 19
          ||| || || || ||||||
Db      60 CATAGCATGACTCCCTGGA 42


RESULT 13
US-09-897-778-370/c
; Sequence 370, Application US/09897778
; Patent No. US20020147143A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, Tongtong
;  APPLICANT:  Marnerakis, Margarita
;  APPLICANT:  Fanger, Gary R.
;  APPLICANT:  Vedvick, Thomas S.
;  APPLICANT:  Carter, Darrick
;  APPLICANT:  Watanabe, Yoshihiro
;  APPLICANT:  Henderson, Robert A.
;  APPLICANT:  Peckham, David W.
;  APPLICANT:  Fanger, Neil
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF LUNG CANCER
;  FILE REFERENCE: 210121.455C16
;  CURRENT APPLICATION NUMBER: US/09/897,778
;  CURRENT FILING DATE:  2001-06-28
;  NUMBER OF SEQ ID NOS: 467
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 370
;   LENGTH: 60
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-897-778-370

  Query Match             63.0%;  Score 12.6;  DB 10;  Length 60;
  Best Local Similarity   78.9%;  Pred. No. 2.2e+03;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 CATTGCCTGGCTTCCTGGA 19
          ||| || || || ||||||
Db      60 CATAGCATGACTCCCTGGA 42


RESULT 14
US-09-864-761-33144/c
; Sequence 33144, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 33144
;   LENGTH: 99
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO L78810.1
;   OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 0.94
;   OTHER INFORMATION: EST_HUMAN HIT: BE693933.1, EVALUE 7.80e-02
;   OTHER INFORMATION: NT HIT: L78810.1, EVALUE 3.00e-49
US-09-864-761-33144

  Query Match             63.0%;  Score 12.6;  DB 10;  Length 99;
  Best Local Similarity   78.9%;  Pred. No. 2.3e+03;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 CATTGCCTGGCTTCCTGGA 19
          | |  | ||||||||||||
Db      62 CTTCACTTGGCTTCCTGGA 44


RESULT 15
US-09-294-093B-1610
; Sequence 1610, Application US/09294093B
; Patent No. US20010051335A1
```

```
; GENERAL INFORMATION:
;  APPLICANT:      Lalgudi, Raghunath, V.
;  APPLICANT:  Ito, Laura, Y.
;  APPLICANT:  Sherman, Bradley, K.
;  TITLE OF INVENTION:  POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN TASSEL
;  FILE REFERENCE:  PL-0009 US
;  CURRENT APPLICATION NUMBER: US/09/294,093B
;  CURRENT FILING DATE:  1999-04-16
;  PRIOR APPLICATION NUMBER:  60/082,567
;  PRIOR FILING DATE:  April 21, 1998
;  NUMBER OF SEQ ID NOS:      6207
;  SOFTWARE:      PERL Program
; SEQ ID NO 1610
;   LENGTH: 100
;   TYPE: DNA
;   ORGANISM: Zea mays
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. US20010051335A1 700344692H1
;   NAME/KEY: unsure
;   LOCATION: 47
;   OTHER INFORMATION: a, t, c, g, or other
US-09-294-093B-1610

  Query Match             63.0%;  Score 12.6;  DB 10;  Length 100;
  Best Local Similarity   78.9%;  Pred. No. 2.3e+03;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 ATTGCCTGGCTTCCTGGAG 20
           ||||| | | |||||||  |
Db      51 ATTGCATAGGTTCCTGGTG 69


Search completed: November 23, 2002, 06:42:14
Job time : 17.25 secs
```

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 25, 2002, 09:10:06 ; Search time 755.55 Seconds
(without alignments)
428.707 Million cell updates/sec

Title: US-09-296-264-19
Perfect score: 20
Sequence: 1 cattgcctggcttcctggag 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 357874

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	15.2	76.0	81	14	BQ840126	BQ840126 mah74g07.
	2	14.8	74.0	50	9	AU107427	AU107427 AU107427
	3	14.2	71.0	61	14	H39535	H39535 yl68a08.r1
	4	14.2	71.0	100	14	H45155	H45155 yo01c09.r1
c	5	14	70.0	65	14	T90432	T90432 ye16h01.r1
c	6	14	70.0	72	14	H41166	H41166 yp64g04.s1
	7	13.8	69.0	45	10	AV833283	AV833283 AV833283
	8	13.8	69.0	61	13	BG923266	BG923266 602825221
c	9	13.6	68.0	63	17	AZ628585	AZ628585 1M0480G13
c	10	13.6	68.0	70	9	AI828161	AI828161 wk32g11.x
c	11	13.6	68.0	97	9	AA506628	AA506628 ni16c06.s
	12	13.4	67.0	33	17	AZ443820	AZ443820 1M0238I06
	13	13.4	67.0	76	9	AU076691	AU076691 AU076691
c	14	13.4	67.0	79	9	AA968879	AA968879 op37h11.s
c	15	13.4	67.0	82	17	AZ579313	AZ579313 1M0363P10
	16	13.2	66.0	68	17	AZ599373	AZ599373 1M0414K09
c	17	13.2	66.0	98	10	AW118112	AW118112 xe03d11.x
c	18	13	65.0	71	12	BF023651	BF023651 EtESTea97
c	19	13	65.0	95	9	AA681695	AA681695 vu76g03.r
c	20	13	65.0	97	9	AA168049	AA168049 ms26d02.r
c	21	13	65.0	98	9	AA168044	AA168044 ms26b02.r
c	22	12.8	64.0	32	13	BI830591	BI830591 603073122
	23	12.8	64.0	40	17	AZ664762	AZ664762 1M0545P18
c	24	12.8	64.0	42	13	BI826734	BI826734 603077210
	25	12.8	64.0	48	17	AZ766335	AZ766335 1M0563P13
c	26	12.8	64.0	54	13	BI829842	BI829842 603079922
	27	12.8	64.0	55	9	AA630141	AA630141 ab99d07.s
c	28	12.8	64.0	55	17	AZ464795	AZ464795 1M0274A21
	29	12.8	64.0	59	17	B03179	B03179 cSRL-16f8-u
c	30	12.8	64.0	60	13	BI830865	BI830865 603080983
c	31	12.8	64.0	63	14	D25834	D25834 HUMGS04204
	32	12.8	64.0	64	9	AA464372	AA464372 zx81d05.r
c	33	12.8	64.0	83	17	AZ767210	AZ767210 1M0566E17
	34	12.8	64.0	83	17	BH408450	BH408450 1007002G0
c	35	12.8	64.0	84	13	BI828236	BI828236 603078115
	36	12.8	64.0	84	17	BH408679	BH408679 1007007B0
	37	12.8	64.0	85	14	W30031	W30031 mc24g08.r1
	38	12.8	64.0	93	17	BH416747	BH416747 1007049H0
	39	12.8	64.0	94	9	AA151756	AA151756 zo29e04.r
	40	12.8	64.0	94	17	BH416896	BH416896 1007051C0
c	41	12.8	64.0	95	9	AA121929	AA121929 zm24f11.s
	42	12.8	64.0	95	12	BF588657	BF588657 7i20h07.x
	43	12.8	64.0	97	14	W30071	W30071 mc23b11.r1
	44	12.8	64.0	98	17	BH412599	BH412599 1007028B0
	45	12.8	64.0	99	17	BH411791	BH411791 1007024A1

ALIGNMENTS

RESULT 1
BQ840126
LOCUS BQ840126 81 bp mRNA linear EST 12-AUG-2002
DEFINITION mah74g07.y1 McCarrey Eddy 18 day preleptotene spermatocytes Mus musculus cDNA clone IMAGE:6366516 5', mRNA sequence.
ACCESSION BQ840126
VERSION BQ840126.1 GI:22209535
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 81)
AUTHORS McCarrey,J., Eddy,M., Marra,M., Hillier,L., Clifton,S., Pape,D., Martin,J., Wylie,T., Dante,M., Bowers,Y., Theising,B., Gibbons,M., Ritter,E., Tsagareishvili,R., Ronko,I., Maguire,L., Kennedy,S., Bennett,J., Waterston,R. and Wilson,R.
TITLE NIEHS Mouse
JOURNAL Unpublished (2002)
COMMENT Contact: McCarrey/Eddy NIEHS Mouse
NIEHS Mouse
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:2043948

```
            Seq primer: -40RP from Gibco
            High quality sequence stop: 69.
FEATURES              Location/Qualifiers
     source           1. .81
                      /organism="Mus musculus"
                      /db_xref="taxon:10090"
                      /clone="IMAGE:6366516"
                      /clone_lib="McCarrey Eddy 18 day preleptotene
                      spermatocytes"
                      /sex="male"
                      /tissue_type="18-day preleptotene spermatocytes"
                      /lab_host="DH10B (phage-resistant)"
                      /note="Organ: testis; Vector: pBluescript SK+ (Stratagene
                      ); Site_1: XhoII; Site_2: EcoRI; cDNA oligo dT-primed
                      [5'-(GA)10-ACTAGTCTCGAGTTTTTTTTTTTTT-3'] and directionally
                      cloned using 5' linkers 5'-AATTCGGCACGAG-3' and
                      5'-CTCGTGCCG-3'. Size selection of >400bp material gives
                      average insert size ranging from 1-2 kb. Library was mass
                      excised (from lambda-UniZAP-XR) and resulting
                      single-stranded phagemids were prepped and tranformed into
                      DH10B. Library constructed and donated by J. McCarrey,
                      Ph.D. (Southwest Foundation for Biomedical Research, Dept.
                      of Genetics); excision done by E.M. Eddy, Ph.D. (National
                      Institutes of Health, National Institute of Environmental
                      Health Sciences)."
BASE COUNT       15 a     24 c     18 g     24 t
ORIGIN

  Query Match             76.0%;  Score 15.2;  DB 14;  Length 81;
  Best Local Similarity   85.0%;  Pred. No. 7.1e+03;
  Matches   17;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 CATTGCCTGGCTTCCTGGAG 20
           |||||||||| || ||| ||
Db      59 CATTGCCTGGATTGCTGCAG 78


RESULT 2
AU107427
LOCUS       AU107427                 50 bp    mRNA    linear   EST 30-AUG-2001
DEFINITION  AU107427 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
            LNG15358, mRNA sequence.
ACCESSION   AU107427
VERSION     AU107427.1  GI:13556948
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Suzuki,Y., Taira,H., Tsunoda,T., Mizushima-Sugano,J., Sese,J., Hata
            ,H., Ota,T., Isogai,T., Tanaka,T., Morishita,S., Okubo,K., Sakaki
            ,Y., Nakamura,Y., Suyama,A. and Sugano,S.
  TITLE     Diverse transcriptional initiation revealed by fine, large-scale
            mapping of mRNA start sites
  JOURNAL   EMBO Rep. 2 (5), 388-393 (2001)
  MEDLINE   21270072
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp
            Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano
            ,S. Construction and characterization of a full length-enriched and
            a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).
FEATURES              Location/Qualifiers
     source           1. .50
                      /organism="Homo sapiens"
                      /db_xref="taxon:9606"
                      /clone="LNG15358"
                      /clone_lib="Sugano Homo sapiens cDNA library"
                      /note="Differential display comparison of untreated and
                      dimethylfumarate treated U937 cells"
BASE COUNT        2 a     13 c     17 g     18 t
ORIGIN

  Query Match             74.0%;  Score 14.8;  DB 9;  Length 50;
  Best Local Similarity   88.9%;  Pred. No. 8.8e+03;
  Matches   16;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       3 TTGCCTGGCTTCCTGGAG 20
           ||||  ||||||||||||
Db      18 TTGCTCGGCTTCCTGGAG 35


RESULT 3
H39535
LOCUS       H39535                   61 bp    mRNA    linear   EST 31-JUL-1995
DEFINITION  y168a08.r1 Soares breast 3NbHBst Homo sapiens cDNA clone
            IMAGE:163382 5' similar to SP:A35098 A35098 MHC CLASS III
            HISTOCOMPATIBILITY ANTIGEN HLA-B-ASSOCIATED TRANSCRIPT 3 - ;, mRNA
            sequence.
ACCESSION   H39535
VERSION     H39535.1  GI:915551
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 61)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman
            ,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J.,
            Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston
            ,R., Williamson,A., Wohldmann,P. and Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Insert Size: 1688
            High quality sequence starts: 1
            High quality sequence stops: 1
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Trace considered overall poor quality
            Possible reversed clone: similarity on wrong strand
            Insert Length: 1688   Std Error: 0.00
            Seq primer: M13RP1
            High quality sequence stop: 1.
FEATURES              Location/Qualifiers
     source           1. .61
                      /organism="Homo sapiens"
                      /db_xref="GDB:577940"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:163382"
                      /clone_lib="Soares breast 3NbHBst"
                      /sex="Female"
                      /dev_stage="adult"
                      /lab_host="DH10B (ampicillin resistant)"
                      /note="Organ: breast; Vector: pT7T3D (Pharmacia) with a
                      modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                      strand cDNA was primed with a Not I - oligo(dT) primer [5'
                      TGTTACCAATCTGAAGTGGGAGCGGCCGCCCTTTTTTTTTTTTTTTTT 3'],
                      double-stranded cDNA was ligated to Eco RI adaptors
                      (Pharmacia), digested with Not I and cloned into the Not I
                      and Eco RI sites of a modified pT7T3 vector (Pharmacia).
                      Library went through one round of normalization to a Cot =
                      20. Library constructed by Bento Soares and M.Fatima
                      Bonaldo."
BASE COUNT       16 a     27 c     10 g      7 t      1 others
ORIGIN
```

```
  Query Match             71.0%;  Score 14.2;  DB 14;  Length 61;
  Best Local Similarity   80.0%;  Pred. No. 1.8e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 CATTGCCTGGCTTCCTGGAG 20
           ||||| ||| |||| || ||
Db       9 CATTGACTGACTTCNTGCAG 28


RESULT 4
H45155
LOCUS       H45155                  100 bp    mRNA    linear   EST 31-JUL-1995
DEFINITION  yo01c09.r1 Soares adult brain N2b5HB55Y Homo sapiens cDNA clone
            IMAGE:176656 5' similar to SP:ECR_DROME P34021 ECDYSONE ;, mRNA
            sequence.
ACCESSION   H45155
VERSION     H45155.1  GI:921207
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 100)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman
            ,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J.,
            Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston
            ,R., Williamson,A., Wohldmann,P. and Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Insert Size: 1229
            High quality sequence starts: 1
            High quality sequence stops: 1
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Trace considered overall poor quality
            Possible reversed clone: similarity on wrong strand
            Insert Length: 1229   Std Error: 0.00
            Seq primer: M13RP1
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .100
                     /organism="Homo sapiens"
                     /db_xref="GDB:3838852"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:176656"
                     /clone_lib="Soares adult brain N2b5HB55Y"
                     /sex="Male"
                     /dev_stage="55-year old"
                     /lab_host="DH10B (ampicillin resistant)"
                     /note="Organ: brain; Vector: pT7T3D (Pharmacia) with a
                     modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCGCTTTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was size selected, ligated to Eco RI
                     adapters (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of a modified pT7T3 vector
                     (Pharmacia). Library went through one round of
                     normalization to a Cot = 53. Library constructed by Bento
                     Soares and M.Fatima Bonaldo. The adult brain RNA was
                     provided by Dr. Donald H. Gilden. Tissue was acquired
                     17-18 hours after death which occurred in consequence of a
                     ruptured aortic aneurysm. RNA was prepared from a pool of
                     tissues representing the following areas of the brain:
                     frontal, parietal, temporal and occipital cortex from the
                     left and right hemispheres, subcortical white matter,
                     basal ganglia, thalamus, cerebellum, midbrain, pons and
                     medulla."
BASE COUNT       19 a     28 c     30 g     22 t      1 others
ORIGIN

  Query Match             71.0%;  Score 14.2;  DB 14;  Length 100;
  Best Local Similarity   84.2%;  Pred. No. 2.1e+04;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 ATTGCCTGGCTTCCTGGAG 20
           | ||||||| |||||| ||
Db      61 AGTGCCTGGTTTCCTGCAG 79


RESULT 5
T90432/c
LOCUS       T90432                  65 bp    mRNA    linear   EST 20-MAR-1995
DEFINITION  ye16h01.r1 Stratagene lung (#937210) Homo sapiens cDNA clone
            IMAGE:117937 5' similar to gb:J03890_rna1 PULMONARY
            SURFACTANT-ASSOCIATED PROTEIN C PRECURSOR (HUMAN);, mRNA sequence.
ACCESSION   T90432
VERSION     T90432.1  GI:718945
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 65)
  AUTHORS   Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiapelli,B.,
            Chissoe,S., Dietrich,N., DuBuque,T., Favello,A., Gish,W., Hawkins
            ,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N., Mardis,E., Moore
            ,B., Morris,M., Parsons,J., Prange,C., Rifkin,L., Rohlfing,T.,
            Schellenberg,K., Soares,M.B., Tan,F., Thierry-Meg,J., Trevaskis,E.,
            Underwood,K., Wohldmann,P., Waterston,R., Wilson,R. and Marra,M.
  TITLE     Generation and analysis of 280,000 human expressed sequence tags
  JOURNAL   Genome Res. 6 (9), 807-828 (1996)
  MEDLINE   97044478
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            High quality sequence starts: 1
            High quality sequence stops: 1
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Trace considered overall poor quality
            Seq primer: M13RP1
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .65
                     /organism="Homo sapiens"
                     /db_xref="GDB:486226"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:117937"
                     /clone_lib="Stratagene lung (#937210)"
                     /sex="male"
                     /dev_stage="72 years"
                     /lab_host="SOLR cells (kanamycin resistant)"
                     /note="Organ: lung; Vector: pBluescript SK-; Site_1: EcoRI
                     ; Site_2: XhoI; Cloned unidirectionally.  Primer: Oligo
                     dT. normal lung.  Average insert size: 1.0 kb; Uni-ZAP XR
                     Vector; ~5' adaptor sequence: 5' GAATTCGGCACGAG 3' ~3'
                     adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTTTT 3'"
BASE COUNT       17 a     13 c     22 g      8 t      5 others
ORIGIN

  Query Match             70.0%;  Score 14;  DB 14;  Length 65;
  Best Local Similarity   93.3%;  Pred. No. 2.2e+04;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       3 TTGCCTGGCTTCCTG 17
```

```
          ||||||||||||| |
Db      55 TTGCCTGGCTTCCNG 41


RESULT 6
H41166/c
LOCUS       H41166                    72 bp    mRNA    linear   EST 16-AUG-1995
DEFINITION  yp64g04.s1 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone
            IMAGE:192246 3' similar to gb:M91036_rna2 HEMOGLOBIN GAMMA-A AND
            GAMMA-G CHAINS (HUMAN);, mRNA sequence.
ACCESSION   H41166
VERSION     H41166.1  GI:917218
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 72)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman
            ,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J.,
            Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston
            ,R., Williamson,A., Wohldmann,P. and Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Insert Size: 1136
            High quality sequence starts: 1
            High quality sequence stops: 1
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Trace considered overall poor quality
            Insert Length: 1136   Std Error: 0.00
            Seq primer: Promega -21m13
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .72
                     /organism="Homo sapiens"
                     /db_xref="GDB:3762035"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:192246"
                     /clone_lib="Soares fetal liver spleen 1NFLS"
                     /sex="male"
                     /dev_stage="20 week-post conception fetus"
                     /lab_host="DH10B (ampicillin resistant)"
                     /note="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia)
                     with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
                     1st strand cDNA was primed with a Pac I - oligo(dT) primer
                     [5' AACTGGAAGAATTAATTAAAGATCTTTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Pac I and cloned into the Pac I
                     and Eco RI sites of the modified pT7T3 vector.  Library
                     went through one round of normalization. Library
                     constructed by Bento Soares and M.Fatima Bonaldo."
BASE COUNT       16 a     18 c     24 g      7 t      7 others
ORIGIN

  Query Match             70.0%;  Score 14;  DB 14;  Length 72;
  Best Local Similarity   100.0%;  Pred. No. 2.3e+04;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       6 CCTGGCTTCCTGGA 19
           ||||||||||||||
Db      61 CCTGGCTTCCTGGA 48


RESULT 7
AV833283
LOCUS       AV833283                  45 bp    mRNA    linear   EST 22-JUN-2001
DEFINITION  AV833283 K. Sato unpublished cDNA library: Hordeum vulgare subsp.
            vulgare shoots germination Hordeum vulgare subsp. vulgare cDNA
            clone bags5a20, mRNA sequence.
ACCESSION   AV833283
VERSION     AV833283.1  GI:14525372
KEYWORDS    EST.
SOURCE      Hordeum vulgare subsp. vulgare.
  ORGANISM  Hordeum vulgare subsp. vulgare
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
            ; Triticeae; Hordeum.
REFERENCE   1  (bases 1 to 45)
  AUTHORS   Sato,K.
  TITLE     Barley EST sequencing project in NIG and Okayama Univ
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Kazuhiro Sato
            Research Institute for Bioresources
            Okayama University, Barley Germplasm Center
            Chuo 2-20-1, Kurashiki, Okayama 710-0046, Japan
            Email: kazsato@rib.okayama-u.ac.jp,
            URL:http://www.rib.okayama-u.ac.jp/barley/
            Sato,K., Saisho,D., Takeda,K., Shini,T. and Kohara,Y. Direct
            submission;
            database:http://www.shigen.nig.ac.jp/barley/Barley.html.
FEATURES             Location/Qualifiers
     source          1. .45
                     /organism="Hordeum vulgare subsp. vulgare"
                     /cultivar="Haruna Nijo"
                     /db_xref="taxon:112509"
                     /clone="bags5a20"
                     /clone_lib="K. Sato unpublished cDNA library: Hordeum
                     vulgare subsp. vulgare shoots germination"
                     /tissue_type="shoots"
                     /dev_stage="germination"
BASE COUNT        9 a     12 c     11 g      9 t      4 others
ORIGIN

  Query Match             69.0%;  Score 13.8;  DB 10;  Length 45;
  Best Local Similarity   78.9%;  Pred. No. 2.3e+04;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 ATTGCCTGGCTTCCTGGAG 20
           || ||||| | ||||| ||
Db       5 ATNGCCTGTCNTCCTGAAG 23


RESULT 8
BG923266
LOCUS       BG923266                  61 bp    mRNA    linear   EST 05-JUN-2001
DEFINITION  602825221F1 NCI_CGAP_Mam6 Mus musculus cDNA clone IMAGE:4954061 5',
            mRNA sequence.
ACCESSION   BG923266
VERSION     BG923266.1  GI:14303742
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 61)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Jeffrey Green M.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10914  row: i  column: 06
```

```
          High quality sequence stop: 61.
FEATURES             Location/Qualifiers
     source          1. .61
                     /organism="Mus musculus"
                     /strain="FVB/N"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:4954061"
                     /clone_lib="NCI_CGAP_Mam6"
                     /sex="female, virgin"
                     /tissue_type="infiltrating ductal carcinoma"
                     /dev_stage="5 months"
                     /lab_host="DH10B"
                     /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                     Library constructed by Life Technologies. Investigator
                     providing samples: Jeffrey Green, M.D., NIH"
BASE COUNT        5 a     16 c     33 g      7 t
ORIGIN

  Query Match             69.0%;  Score 13.8;  DB 13;  Length 61;
  Best Local Similarity   88.2%;  Pred. No. 2.6e+04;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       4 TGCCTGGCTTCCTGGAG 20
           ||||||| | |||||||
Db      18 TGCCTGGGTGCCTGGAG 34


RESULT 9
AZ628585/c
LOCUS       AZ628585                 63 bp    DNA     linear   GSS 13-DEC-2000
DEFINITION  1M0480G13R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0480G13 R, DNA sequence.
ACCESSION   AZ628585
VERSION     AZ628585.1  GI:11750775
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 63)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000  Std Error: 0.00
            Plate: 0480  row: G  column: 13
            Seq primer: CACACAGGAAACAGCTATGACC
            Class: plasmid ends
            High quality sequence stop: 63.
FEATURES             Location/Qualifiers
     source          1. .63
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC1M0480G13"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
BASE COUNT       14 a     20 c     14 g     15 t
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 17;  Length 63;
  Best Local Similarity   80.0%;  Pred. No. 3.3e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 CATTGCCTGGCTTCCTGGAG 20
           |||||||||   | ||||||
Db      42 CATTGCCTGCAGTTCTGGAG 23


RESULT 10
AI828161/c
LOCUS       AI828161                 70 bp    mRNA    linear   EST 26-AUG-1999
DEFINITION  wk32g11.x1 NCI_CGAP_Brn25 Homo sapiens cDNA clone IMAGE:2414084 3'
            similar to SW:DYJ2_HUMAN O43237 DYNEIN LIGHT INTERMEDIATE CHAIN 2,
            CYTOSOLIC ;, mRNA sequence.
ACCESSION   AI828161
VERSION     AI828161.1  GI:5448832
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 70)
  AUTHORS   NCI/NINDS-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute / National Institute of Neurological
            Disorders and Stroke, Brain Tumor Genome Anatomy Project
            (CGAP/BTGAP), Tumor Gene Index
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D.,
            Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
            Bonaldo, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html

            Trace considered overall poor quality
            Seq primer: -40UP from Gibco
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .70
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2414084"
                     /clone_lib="NCI_CGAP_Brn25"
                     /tissue_type="anaplastic oligodendroglioma"
                     /lab_host="DH10B"
                     /note="Organ: brain; Vector: pT7T3D-Pac (Pharmacia) with a
                     modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer [5'
```

TGTTACCAATCTGAAGTGGGAGCGGCCGCATAGGTTTTTTTTTTTTTTTTTTTTTTTTT
T 3']; double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified pT7T3 vector.
Library is normalized, and was constructed by Bento
Soares and M.Fatima Bonaldo."
BASE COUNT 22 a 13 c 17 g 18 t
ORIGIN

Query Match 68.0%; Score 13.6; DB 9; Length 70;
Best Local Similarity 80.0%; Pred. No. 3.4e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
Qy      1 CATTGCCTGGCTTCCTGGAG 20
          ||||||| | ||| || |||
Db     31 CATTGCCCGTCTTTCTAGAG 12
```

RESULT 11
AA506628/c
LOCUS AA506628 97 bp mRNA linear EST 20-AUG-1997
DEFINITION n116c06.s1 NCI_CGAP_Co4 Homo sapiens cDNA clone IMAGE:968170 3' similar to gb:L02426 26S PROTEASE SUBUNIT 4 (HUMAN);, mRNA sequence.
ACCESSION AA506628
VERSION AA506628.1 GI:2242775
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 97)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Elias Campo, M.D., Michael R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html

Trace considered overall poor quality
Insert Length: 2226 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 1.

FEATURES Location/Qualifiers
source 1. .97
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:968170"
/clone_lib="NCI_CGAP_Co4"
/sex="pooled"
/tissue_type="colon"
/lab_host="DH10B"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; 1st strand cDNA was prepared from pooled colon tumor tissue, and was then primed with a Not I - oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. This library is not normalized. Library constructed by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT 30 a 28 c 23 g 16 t
ORIGIN

Query Match 68.0%; Score 13.6; DB 9; Length 97;
Best Local Similarity 80.0%; Pred. No. 3.9e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
Qy      1 CATTGCCTGGCTTCCTGGAG 20
          ||||| |||| |||||| |
Db     62 CATTGGATGGCCTCCTGGTG 43
```

RESULT 12
AZ443820
LOCUS AZ443820 33 bp DNA linear GSS 04-OCT-2000
DEFINITION 1M0238I06F Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC1M0238I06 F, DNA sequence.
ACCESSION AZ443820
VERSION AZ443820.1 GI:10592178
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 33)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D.,Weiss,R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0238 row: I column: 06
Seq primer: CGTTGTAAAACGACGGCCAGT
Class: plasmid ends
High quality sequence stop: 33.

FEATURES Location/Qualifiers
source 1. .33
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0238I06"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptored DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptored mouse DNA was annealed to adaptored vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."
BASE COUNT 4 a 11 c 8 g 10 t
ORIGIN

Query Match 67.0%; Score 13.4; DB 17; Length 33;
Best Local Similarity 93.3%; Pred. No. 3.1e+04;

```
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 CATTGCCTGGCTTCC 15
          ||||||||||||  ||
Db     10 CATTGCCTGGCTGCC 24


RESULT 13
AU076691
LOCUS       AU076691                 76 bp    mRNA    linear   EST 04-MAY-2000
DEFINITION  AU076691 Sugano cDNA library Homo sapiens cDNA clone HEP03550
            similar to 5'-end region of Human mRNA encoding prothrombin, mRNA
            sequence.
ACCESSION   AU076691
VERSION     AU076691.1  GI:7439179
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 76)
  AUTHORS   Suzuki,Y., Ishihara,D., Sasaki,M., Nakagawa,H., Hata,H., Tsunoda,T.
            , Watanabe,M., Komatsu,T., Ota,T., Isogai,T., Suyama,A. and Sugano
            ,S.
  TITLE     Statistical analysis of the 5' untranslated region of human mRNA
            using 'Oligo-Capped' cDNA libraries
  JOURNAL   Genomics 64 (3), 286-297 (2000)
  MEDLINE   20221373
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp
            Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano
            ,S. Construction and characterization of a full length-enriched and
            a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997)
             This clone was obtained from a 'full length-enriched' cDNA library
            constructed by 'Oligo-Capping' method. The coding region starts
            from the 50 bp upstream to the 3'-end.
FEATURES             Location/Qualifiers
     source          1. .76
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="HEP03550"
                     /clone_lib="Sugano cDNA library"
                     /note="The cDNA was prepared using the anchor primer,
                     H-T11G, from Genhunter"
BASE COUNT       10 a     27 c     24 g     15 t
ORIGIN

  Query Match          67.0%;  Score 13.4;  DB 9;  Length 76;
  Best Local Similarity   93.3%;  Pred. No. 4.3e+04;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      4 TGCCTGGCTTCCTGG 18
          |||||||||  |||||
Db     52 TGCCTGGCTGCCTGG 66


RESULT 14
AA968879/c
LOCUS       AA968879                 79 bp    mRNA    linear   EST 07-JUL-1998
DEFINITION  op37h11.s1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone
            IMAGE:1579077 3', mRNA sequence.
ACCESSION   AA968879
VERSION     AA968879.1  GI:3144059
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 79)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 634   Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 58.
FEATURES             Location/Qualifiers
     source          1. .79
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1579077"
                     /clone_lib="Soares_NFL_T_GBC_S1"
                     /lab_host="DH10B"
                     /note="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with
                     a modified polylinker; Site_1: Not I; Site_2: Eco RI;
                     Equal amounts of plasmid DNA from three normalized
                     libraries (fetal lung NbHL19W, testis NHT, and B-cell
                     NCI_CGAP_GCB1) were mixed, and ss circles were made in
                     vitro. Following HAP purification, this DNA was used as
                     tracer in a subtractive hybridization reaction. The driver
                     was PCR-amplified cDNAs from pools of 5,000 clones made
                     from the same 3 libraries. The pools consisted of
                     I.M.A.G.E. clones 297480-302087, 682632-687239,
                     726408-728711, and 729096-731399. Subtraction by Bento
                     Soares and M. Fatima Bonaldo.  "
BASE COUNT       15 a     16 c     23 g     24 t      1 others
ORIGIN

  Query Match          67.0%;  Score 13.4;  DB 9;  Length 79;
  Best Local Similarity   93.3%;  Pred. No. 4.4e+04;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      6 CCTGGCTTCCTGGAG 20
          ||||||||||||  ||
Db     42 CCTGGCTTCCTGCAG 28


RESULT 15
AZ579313/c
LOCUS       AZ579313                 82 bp    DNA     linear   GSS 13-DEC-2000
DEFINITION  1M0363P10R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0363P10 R, DNA sequence.
ACCESSION   AZ579313
VERSION     AZ579313.1  GI:11693742
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 82)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0363  row: P  column: 10
            Seq primer: CACACAGGAAACAGCTATGACC
            Class: plasmid ends
```

```
          High quality sequence stop: 82.
FEATURES             Location/Qualifiers
     source          1. .82
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC1M0363P10"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
BASE COUNT       14 a     27 c     26 g     15 t
ORIGIN

  Query Match             67.0%;  Score 13.4;  DB 17;  Length 82;
  Best Local Similarity   93.3%;  Pred. No. 4.5e+04;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       4 TGCCTGGCTTCCTGG 18
           |||| ||||||||||
Db      67 TGCCCGGCTTCCTGG 53


Search completed: November 26, 2002, 04:09:29
Job time : 765.8 secs
```

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 3, 2002, 14:46:40 ; Search time 302.2 Seconds
(without alignments)
1926.063 Million cell updates/sec

Title: US-09-296-264-20
Perfect score: 20
Sequence: 1 cccagggcactcatggctat 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 995600

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
c	1	18.6	93.0	21	6	AX153915	AX153915 Sequence
	2	14	70.0	31	6	AX248874	AX248874 Sequence
c	3	13.8	69.0	24	6	AX289345	AX289345 Sequence
	4	13.6	68.0	87	6	A57087	A57087 Sequence 40
	5	13.6	68.0	87	6	A80345	A80345 Sequence 40
	6	13.6	68.0	96	10	AY041968	AY041968 Oecomys b
	7	13.4	67.0	78	6	AX404028	AX404028 Sequence
	8	13.4	67.0	93	6	E29108	E29108 Expression
	9	13.2	66.0	24	6	AR091297	AR091297 Sequence
c	10	13.2	66.0	24	6	AR104159	AR104159 Sequence
	11	13.2	66.0	24	6	AR137679	AR137679 Sequence
	12	13.2	66.0	61	6	E29076	E29076 Modified in
	13	13.2	66.0	67	9	S75425S04	S75432 collagen ty
	14	13.2	66.0	89	10	AY042047	AY042047 Sigmodon
	15	13.2	66.0	94	10	AY041953	AY041953 Oryzomys
c	16	12.8	64.0	20	6	AX293978	AX293978 Sequence
c	17	12.8	64.0	51	6	AX157975	AX157975 Sequence
	18	12.8	64.0	57	6	AR034237	AR034237 Sequence
	19	12.8	64.0	57	6	AR156699	AR156699 Sequence
	20	12.8	64.0	57	6	I17650	I17650 Sequence 5
c	21	12.8	64.0	66	6	E03396	E03396 DNA encodin
c	22	12.8	64.0	69	6	I66412	I66412 Sequence 10
	23	12.8	64.0	74	9	HSACAA3	X65141 H.sapiens A
c	24	12.8	64.0	99	10	AF357349	AF357349 Mus muscu
c	25	12.6	63.0	21	6	AX096010	AX096010 Sequence
	26	12.6	63.0	23	6	AR089639	AR089639 Sequence
	27	12.6	63.0	23	6	AR118873	AR118873 Sequence
c	28	12.6	63.0	29	6	AR106731	AR106731 Sequence
c	29	12.6	63.0	36	6	A00470	A00470 Nucleotide
c	30	12.6	63.0	36	6	I04603	I04603 Sequence 3
c	31	12.6	63.0	36	6	I05147	I05147 Sequence 4
	32	12.6	63.0	57	6	AX404702	AX404702 Sequence
c	33	12.6	63.0	69	6	AR171551	AR171551 Sequence
c	34	12.6	63.0	69	6	BD005571	BD005571 Compositi
	35	12.6	63.0	77	10	AY042028	AY042028 Sigmodon
c	36	12.6	63.0	78	1	SATRNMET	X01223 Sulfolobus
c	37	12.6	63.0	89	6	A69886	A69886 Sequence 15
	38	12.6	63.0	92	9	HCOL4A6S37	U46995 Human type
	39	12.6	63.0	93	6	AX404703	AX404703 Sequence
c	40	12.6	63.0	96	6	AR165711	AR165711 Sequence
	41	12.6	63.0	96	9	HUMCG3A1N	M11134 Human proco
	42	12.6	63.0	98	10	AY041785	AY041785 Holochilu
c	43	12.4	62.0	18	6	AR169374	AR169374 Sequence
	44	12.4	62.0	18	6	I25834	I25834 Sequence 14
c	45	12.4	62.0	20	6	AX293987	AX293987 Sequence

ALIGNMENTS

RESULT 1
AX153915/c
LOCUS AX153915 21 bp DNA linear PAT 22-JUN-2001
DEFINITION Sequence 13 from Patent WO0138576.
ACCESSION AX153915
VERSION AX153915.1 GI:14535529
KEYWORDS .
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 21)
AUTHORS Cargill,M., Ireland,J.S. and Lander,E.S.
TITLE Human single nucleotide polymorphisms
JOURNAL Patent: WO 0138576-A 13 31-MAY-2001;

```
          WHITEHEAD INSTITUTE FOR BIOMEDICAL RESEARCH (US)
FEATURES             Location/Qualifiers
     source          1. .21
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT        3 a      5 c      9 g      3 t      1 others
ORIGIN

  Query Match            93.0%;  Score 18.6;  DB 6;  Length 21;
  Best Local Similarity  94.7%;  Pred. No. 50;
  Matches   18;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CCCAGGGCACTCATGGCTA 19
           |||||||| ||||||||||
Db      19 CCCAGGGCRCTCATGGCTA 1


RESULT 2
AX248874
LOCUS       AX248874                 31 bp    DNA     linear   PAT 28-SEP-2001
DEFINITION  Sequence 953 from Patent WO0166800.
ACCESSION   AX248874
VERSION     AX248874.1  GI:15863497
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 31)
  AUTHORS   Cargill,M., Ireland,J.S. and Lander,E.S.
  TITLE     Human single nucleotide polymorphisms
  JOURNAL   Patent: WO 0166800-A 953 13-SEP-2001;
            WHITEHEAD INSTITUTE FOR BIOMEDICAL RESEARCH (US)
FEATURES             Location/Qualifiers
     source          1. .31
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT        7 a      6 c     11 g      6 t      1 others
ORIGIN

  Query Match            70.0%;  Score 14;  DB 6;  Length 31;
  Best Local Similarity  87.5%;  Pred. No. 1.5e+04;
  Matches   14;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       5 GGGCACTCATGGCTAT 20
           |||| |||||||:|||
Db       4 GGGCCCTCATGGMTAT 19


RESULT 3
AX289345/c
LOCUS       AX289345                 24 bp    DNA     linear   PAT 21-NOV-2001
DEFINITION  Sequence 1107 from Patent WO0179548.
ACCESSION   AX289345
VERSION     AX289345.1  GI:17051028
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Barany,F., Zirvi,M., Gerry,N.P., Favis,R. and Kliman,R.
  TITLE     Method of designing addressable array for detection of nucleic acid
            sequence differences using ligase detection reaction
  JOURNAL   Patent: WO 0179548-A 1107 25-OCT-2001;
            CORNELL RESEARCH FOUNDATION, INC. (US)
FEATURES             Location/Qualifiers
     source          1. .24
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="Hypothetical Probe Sequence"
BASE COUNT        3 a      7 c      7 g      7 t
ORIGIN

  Query Match            69.0%;  Score 13.8;  DB 6;  Length 24;
  Best Local Similarity  88.2%;  Pred. No. 1.9e+04;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       3 CAGGGCACTCATGGCTA 19
           || |||||||| |||||
Db      20 CAAGGCACTCAAGGCTA 4


RESULT 4
A57087
LOCUS       A57087                   87 bp    DNA     linear   PAT 03-MAR-1998
DEFINITION  Sequence 40 from Patent WO9628552.
ACCESSION   A57087
VERSION     A57087.1  GI:3713070
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 87)
  AUTHORS   Perron,H., Mandrand,B., Mallet,F., Bedin,F. and Beseme,F.
  TITLE     MSRV-1 virus and pathogenic and/or infectious agent MSRV-2
            associated with rheumatoid arthritis
  JOURNAL   Patent: WO 9628552-A 40 19-SEP-1996;
            BIO MERIEUX (FR)
COMMENT     Other publication JP 8322579 961210
            Other publication NO 964760 961108
            Other publication AU 5007396 961002
            Other publication FR 2731356 960913
            Other publication CA 2171242 960910.
FEATURES             Location/Qualifiers
     source          1. .87
                     /organism="unidentified"
                     /db_xref="taxon:32644"
BASE COUNT       11 a     30 c     27 g     19 t
ORIGIN

  Query Match            68.0%;  Score 13.6;  DB 6;  Length 87;
  Best Local Similarity  80.0%;  Pred. No. 2.3e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 CCCAGGGCACTCATGGCTAT 20
           |||||||| || | | ||||
Db       8 CCCAGGGCGCTGAAGCCTAT 27


RESULT 5
A80345
LOCUS       A80345                   87 bp    DNA     linear   PAT 20-OCT-1999
DEFINITION  Sequence 40 from Patent EP0731168.
ACCESSION   A80345
VERSION     A80345.1  GI:6093072
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 87)
  AUTHORS   Perron,H. and Bedin,F.
  TITLE     MSRV-1 VIRUS AND PATHOGENIC AND/OR INFECTIOUS AGENT MSRV-2
            ASSOCIATED WITH RHEUMATOID ARTHRITIS
  JOURNAL   Patent: EP 0731168-A 40 11-SEP-1996;
            BIO MERIEUX (FR)
FEATURES             Location/Qualifiers
     source          1. .87
                     /organism="unidentified"
                     /db_xref="taxon:32644"
BASE COUNT       11 a     30 c     27 g     19 t
ORIGIN

  Query Match            68.0%;  Score 13.6;  DB 6;  Length 87;
  Best Local Similarity  80.0%;  Pred. No. 2.3e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;
```

```
Qy       1 CCCAGGGCACTCATGGCTAT 20
           |||||||| || | | ||||
Db       8 CCCAGGGCGCTGAAGCCTAT 27


RESULT 6
AY041968
LOCUS       AY041968                 96 bp    DNA     linear   ROD 30-JUN-2002
DEFINITION  Oecomys bicolor clone Obic301 B1 SINE retrotransposon.
ACCESSION   AY041968
VERSION     AY041968.1  GI:21632475
KEYWORDS    .
SOURCE      bicolored arboreal rice rat.
  ORGANISM  Oecomys bicolor
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Sigmodontinae;
            Oecomys.
REFERENCE   1  (bases 1 to 96)
  AUTHORS   Rinehart,T.A., Grahn,R.A. and Wichman,H.A.
  TITLE     Coordinated LINE and SINE activity in rodents supports a shared
            retrotransposition mechanism
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 96)
  AUTHORS   Rinehart,T.A., Grahn,R.A. and Wichman,H.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-JUN-2001) Department of Biological Sciences,
            University of Idaho, Gibb Hall, 443051, Moscow, ID 83844-3051, USA
FEATURES             Location/Qualifiers
     source          1. .96
                     /organism="Oecomys bicolor"
                     /specimen_voucher="University of New Mexico, NK12701"
                     /db_xref="taxon:48011"
                     /clone="Obic301"
     repeat_region   1. .96
                     /note="B1 SINE retrotransposon"
                     /rpt_type=dispersed
BASE COUNT       26 a     25 c     25 g     20 t
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 10;  Length 96;
  Best Local Similarity   80.0%;  Pred. No. 2.3e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 CCCAGGGCACTCATGGCTAT 20
           |||||| ||| || | ||||
Db      75 CCCAGGACACCCAAGCCTAT 94


RESULT 7
AX404028
LOCUS       AX404028                 78 bp    DNA     linear   PAT 14-JUN-2002
DEFINITION  Sequence 25 from Patent EP1195161.
ACCESSION   AX404028
VERSION     AX404028.1  GI:21437354
KEYWORDS    .
SOURCE      dog.
  ORGANISM  Canis familiaris
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
REFERENCE   1
  AUTHORS   Morsey,M.A., Sheppard,M.G. and Wheeler,D.W.
  TITLE     Anti-ige vaccines
  JOURNAL   Patent: EP 1195161-A 25 10-APR-2002;
            Pfizer Products Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .78
                     /organism="Canis familiaris"
                     /db_xref="taxon:9615"
BASE COUNT       12 a     36 c     22 g      8 t
ORIGIN

  Query Match             67.0%;  Score 13.4;  DB 6;  Length 78;
  Best Local Similarity   93.3%;  Pred. No. 3e+04;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 CCCAGGGCACTCATG 15
           |||||||| ||||||
Db      25 CCCAGGGCCCTCATG 39


RESULT 8
E29108
LOCUS       E29108                   93 bp    DNA     linear   PAT 18-JUN-2001
DEFINITION  Expression regulatory region of stress protein HSP47 and
            utilization thereof.
ACCESSION   E29108
VERSION     E29108.1  GI:13020964
KEYWORDS    JP 1999243968-A/1.
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 93)
  AUTHORS   Kazuhiro,N.
  TITLE     Expression regulatory region of stress protein HSP47 and
            utilization thereof
  JOURNAL   Patent: JP 1999243968-A 1 14-SEP-1999;
            SCIENCE & TECH AGENCY
COMMENT     OS   Unidentified
            PN   JP 1999243968-A/1
            PD   14-SEP-1999
            PF   05-MAR-1998 JP 1998071489
            PR
            PI   KAZUHIRO NAGATA
            PC   C12N15/09,A61K31/70,A61K48/00//(C12N15/09,C12R1:91),C12N15/00,
            PC   (C12N15/00,C12R1:91)
            CC   Strandedness: Double;
            CC   Topology: Linear;
            FH   Key             Location/Qualifiers
            FT   source          1. .93
            FT                   /organism='Unidentified'.
FEATURES             Location/Qualifiers
     source          1. .93
                     /organism="unidentified"
                     /db_xref="taxon:32644"
BASE COUNT       18 a     26 c     29 g     20 t
ORIGIN

  Query Match             67.0%;  Score 13.4;  DB 6;  Length 93;
  Best Local Similarity   93.3%;  Pred. No. 3e+04;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 CCAGGGCACTCATGG 16
           ||||||||| |||||
Db      11 CCAGGGCACGCATGG 25


RESULT 9
AR091297
LOCUS       AR091297                 24 bp    DNA     linear   PAT 07-SEP-2000
DEFINITION  Sequence 10 from patent US 5994094.
ACCESSION   AR091297
VERSION     AR091297.1  GI:10018052
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 24)
  AUTHORS   Hotten,G., Neidhardt,H. and Paulista,M.
  TITLE     Growth/differentiation factor of the TGF-.beta. family
  JOURNAL   Patent: US 5994094-A 10 30-NOV-1999;
FEATURES             Location/Qualifiers
     source          1. .24
                     /organism="unknown"
BASE COUNT        8 a      7 c      5 g      4 t
ORIGIN
```

```
  Query Match             66.0%;  Score 13.2;  DB 6;  Length 24;
  Best Local Similarity   83.3%;  Pred. No. 4e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 CCAGGGCACTCATGGCTA 19
          |||||||||| ||| | |
Db      2 CCAGGGCACTAATGTCAA 19


RESULT 10
AR104159/c
LOCUS       AR104159                  24 bp    DNA     linear   PAT 14-FEB-2001
DEFINITION  Sequence 8 from patent US 6093542.
ACCESSION   AR104159
VERSION     AR104159.1  GI:12816867
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 24)
  AUTHORS   Romano,J., Pal,R. and Shurtliff,R.
  TITLE     Isothermal transcription based amplification assay for the
            detection and quantitation of macrophage derived chemokine RNA
  JOURNAL   Patent: US 6093542-A 8 25-JUL-2000;
FEATURES             Location/Qualifiers
     source          1. .24
                     /organism="unknown"
BASE COUNT        4 a      8 c      8 g      4 t
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 6;  Length 24;
  Best Local Similarity   83.3%;  Pred. No. 4e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 CCCAGGGCACTCATGGCT 18
          ||||||||||||  || |
Db     21 CCCAGGGCACTCTGGGAT 4


RESULT 11
AR137679
LOCUS       AR137679                  24 bp    DNA     linear   PAT 16-JUN-2001
DEFINITION  Sequence 20 from patent US 6197550.
ACCESSION   AR137679
VERSION     AR137679.1  GI:14479188
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 24)
  AUTHORS   Hotten,G., Neidhardt,H., Bechtold,R. and Pohl,J.
  TITLE     DNA sequences encoding growth/differentiation
  JOURNAL   Patent: US 6197550-A 20 06-MAR-2001;
FEATURES             Location/Qualifiers
     source          1. .24
                     /organism="unknown"
BASE COUNT        8 a      7 c      5 g      4 t
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 6;  Length 24;
  Best Local Similarity   83.3%;  Pred. No. 4e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 CCAGGGCACTCATGGCTA 19
          |||||||||| ||| | |
Db      2 CCAGGGCACTAATGTCAA 19


RESULT 12
E29076
LOCUS       E29076                    61 bp    DNA     linear   PAT 18-JUN-2001
DEFINITION  Modified interferon tau-3.
ACCESSION   E29076
VERSION     E29076.1  GI:13025524
KEYWORDS    JP 1999042089-A/11.
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 61)
  AUTHORS   Masako,I. and Takashi,N.
  TITLE     Modified interferon tau-3
  JOURNAL   Patent: JP 1999042089-A 11 16-FEB-1999;
            SANKYO CO LTD
COMMENT     OS   Unidentified
            PN   JP 1999042089-A/11
            PD   16-FEB-1999
            PF   29-JUL-1997 JP 1997203137
            PR
            PI   MASAKO ISHIMURA,TAKASHI NISHIGAKI
            PC   C12N15/09,A61K38/21,C07K7/06,C07K14/555,C12N1/21,C12P21/02//
            PC   (C12N15/09,C12R1:91),(C12P21/02,C12R1:19),C12N15/00,A61K37/66,
            PC   (C12N15/00,C12R1:91)
            CC   Strandedness: Single;
            CC   Topology: Linear;
            FH   Key             Location/Qualifiers
            FT   source          1. .61
            FT                   /organism='Unidentified'.
FEATURES             Location/Qualifiers
     source          1. .61
                     /organism="unidentified"
                     /db_xref="taxon:32644"
BASE COUNT       16 a     18 c     16 g     11 t
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 6;  Length 61;
  Best Local Similarity   83.3%;  Pred. No. 3.9e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 CCAGGGCACTCATGGCTA 19
          ||||||||  |||| |||
Db     36 CCAGGGCATCCATGTCTA 53


RESULT 13
S75425S04
LOCUS       S75425S04                 67 bp    DNA     linear   PRI 02-JUN-2000
DEFINITION  collagen type VI alpha 2(VI) chain {exons 1-19 and intron/exon
            junctions} [human, Genomic, 67 nt, segment 4 of 19].
ACCESSION   S75432
VERSION     S75432.1  GI:241989
KEYWORDS    .
SEGMENT     4 of 19
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 67)
  AUTHORS   Saitta,B., Wang,Y.M., Renkart,L., Zhang,R.Z., Pan,T.C., Timpl,R.
            and Chu,M.L.
  TITLE     The exon organization of the triple-helical coding regions of the
            human alpha 1(VI) and alpha 2(VI) collagen genes is highly similar
  JOURNAL   Genomics 11 (1), 145-153 (1991)
  MEDLINE   92112205
   PUBMED   1765372
  REMARK    GenBank staff at the National Library of Medicine created this
            entry [NCBI gibbsq 75432] from the original journal article.
            This sequence comes from Figure 4.
            Map location: chromosome 21.
FEATURES             Location/Qualifiers
     source          1. .67
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="21"
     exon            31. .57
                     /number=4
BASE COUNT       12 a     23 c     17 g     15 t
```

```
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 9;  Length 67;
  Best Local Similarity   83.3%;  Pred. No. 3.9e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 CCCAGGGCACTCATGGCT 18
          ||||||||  || |||||
Db     26 CCCAGGGCGTTCCTGGCT 43


RESULT 14
AY042047
LOCUS       AY042047                  89 bp    DNA     linear   ROD 30-JUN-2002
DEFINITION  Sigmodon mascotensis clone Smas171 B1 SINE retrotransposon.
ACCESSION   AY042047
VERSION     AY042047.1  GI:21632554
KEYWORDS    .
SOURCE      Jaliscan cotton rat.
  ORGANISM  Sigmodon mascotensis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Sigmodontinae;
            Sigmodon.
REFERENCE   1  (bases 1 to 89)
  AUTHORS   Rinehart,T.A., Grahn,R.A. and Wichman,H.A.
  TITLE     Coordinated LINE and SINE activity in rodents supports a shared
            retrotransposition mechanism
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 89)
  AUTHORS   Rinehart,T.A., Grahn,R.A. and Wichman,H.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-JUN-2001) Department of Biological Sciences,
            University of Idaho, Gibb Hall, 443051, Moscow, ID 83844-3051, USA
FEATURES             Location/Qualifiers
     source          1. .89
                     /organism="Sigmodon mascotensis"
                     /specimen_voucher="University of Idaho, JS2014"
                     /db_xref="taxon:42416"
                     /clone="Smas171"
     repeat_region   1. .89
                     /note="B1 SINE retrotransposon"
                     /rpt_type=dispersed
BASE COUNT       24 a     25 c     22 g     18 t
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 10;  Length 89;
  Best Local Similarity   83.3%;  Pred. No. 3.8e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 CCAGGGCACTCATGGCTA 19
          ||||||||  || |||||
Db     67 CCAGGGCAGCCAGGGCTA 84


RESULT 15
AY041953
LOCUS       AY041953                  94 bp    DNA     linear   ROD 30-JUN-2002
DEFINITION  Oryzomys albigularis clone Oalb2m B1 SINE retrotransposon.
ACCESSION   AY041953
VERSION     AY041953.1  GI:21632460
KEYWORDS    .
SOURCE      Tomes's rice rat.
  ORGANISM  Oryzomys albigularis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Sigmodontinae;
            Oryzomys.
REFERENCE   1  (bases 1 to 94)
  AUTHORS   Rinehart,T.A., Grahn,R.A. and Wichman,H.A.
  TITLE     Coordinated LINE and SINE activity in rodents supports a shared
            retrotransposition mechanism
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 94)
  AUTHORS   Rinehart,T.A., Grahn,R.A. and Wichman,H.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-JUN-2001) Department of Biological Sciences,
            University of Idaho, Gibb Hall, 443051, Moscow, ID 83844-3051, USA
FEATURES             Location/Qualifiers
     source          1. .94
                     /organism="Oryzomys albigularis"
                     /specimen_voucher="TCWC at Texas A&M University, MUR17"
                     /db_xref="taxon:56215"
                     /clone="Oalb2m"
     repeat_region   1. .94
                     /note="B1 SINE retrotransposon"
                     /rpt_type=dispersed
BASE COUNT       25 a     23 c     29 g     17 t
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 10;  Length 94;
  Best Local Similarity   83.3%;  Pred. No. 3.8e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 CCAGGGCACTCATGGCTA 19
          ||||| || ||| |||||
Db     72 CCAGGACAGTCAGGGCTA 89


Search completed: December  3, 2002, 18:14:13
Job time : 308.2 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 05:52:31 ; Search time 98.55 Seconds
(without alignments)
457.027 Million cell updates/sec

Title: US-09-296-264-20
Perfect score: 20
Sequence: 1 cccagggcactcatggctat 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 2390332

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	20	100.0	20	21	AAZ31450	Human neuropilin m
c	2	17.4	87.0	21	22	AAH62112	Neuropilin 1 (NRP1
	3	16	80.0	20	21	AAZ31441	Human neuropilin m
c	4	14.4	72.0	39	24	ABK88692	Human transmembran
c	5	14.4	72.0	39	24	ABK13410	Tumour necrosis fa
	6	14.2	71.0	23	17	AAT43862	Mouse ob gene reve
	7	14.2	71.0	51	22	AAL27305	Human SNP oligonuc
	8	14.2	71.0	51	22	AAL27306	Human SNP oligonuc
	9	14.2	71.0	51	22	AAL27308	Human SNP oligonuc
c	10	14.2	71.0	60	24	ABN39029	Human spliced tran
c	11	14	70.0	21	22	AAF74885	Chicken neuropilin
	12	14	70.0	42	22	AAF74884	Chicken neuropilin
	13	13.8	69.0	20	24	ABQ61256	Human aquaporin 5
c	14	13.8	69.0	24	24	ABI84610	Capture oligonucle
	15	13.8	69.0	24	24	ABI84611	Capture oligonucle
	16	13.8	69.0	95	19	AAX11689	Human biallelic po
	17	13.8	69.0	95	19	AAX12861	Human biallelic po
	18	13.6	68.0	60	24	ABN32461	Human spliced tran
	19	13.6	68.0	60	24	ABN50481	Human spliced tran
	20	13.6	68.0	87	17	AAT38005	Clone MSRV1polPR i
c	21	13.6	68.0	88	22	ABA69407	Human foetal liver
c	22	13.6	68.0	88	22	ABA36343	Probe #14809 for g
c	23	13.6	68.0	88	24	ABS17620	Human genome-deriv
	24	13.4	67.0	93	20	AAZ27697	Intron from HSP47
	25	13.2	66.0	24	14	AAQ47723	TGF-beta-like clon
	26	13.2	66.0	24	16	AAQ83703	TGF-beta (MP-52) p
c	27	13.2	66.0	24	20	AAX90284	Macrophage derived
	28	13.2	66.0	33	20	AAX59531	PCR primer used to
	29	13.2	66.0	61	20	AAX18466	PCR primer tau3-9
c	30	13	65.0	18	14	AAQ40337	p58 nested primer
c	31	13	65.0	21	19	AAZ26653	Human polymorphic
c	32	13	65.0	21	22	AAF96423	Human gene single
c	33	12.8	64.0	20	20	AAX58467	PCR primer P38-S25
c	34	12.8	64.0	20	24	ABI94020	Capture oligonucle
c	35	12.8	64.0	25	21	AAA68775	Bacteriophage 96 O
	36	12.8	64.0	25	22	AAS43234	Human Oestrogen re
c	37	12.8	64.0	29	21	AAA04666	Polymorphic fragme
c	38	12.8	64.0	31	16	AAQ85848	MHC class I allele
c	39	12.8	64.0	31	16	AAQ85778	MHC CW3 amplificat
c	40	12.8	64.0	31	16	AAQ85742	MHC CW3 cDNA nts 1
	41	12.8	64.0	31	22	AAI30465	Human single nucle
c	42	12.8	64.0	50	19	AAV62498	Plasmid p2ZOp2J-3
c	43	12.8	64.0	50	22	AAL28077	Human SNP oligonuc
	44	12.8	64.0	51	22	AAL27307	Human SNP oligonuc
c	45	12.8	64.0	51	22	AAL28078	Human SNP oligonuc

ALIGNMENTS

RESULT 1
AAZ31450
ID AAZ31450 standard; DNA; 20 BP.
XX
AC AAZ31450;
XX
DT 07-FEB-2000 (first entry)
XX
DE Human neuropilin mRNA specific antisense oligo GTI3621.
XX
KW Neuropilin; human; growth; metastasis; tumor; neovascularisation;
KW cancer; papilloma; diabetic retinopathy; antisense; ss.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9955855-A2.
XX
PD 04-NOV-1999.
XX
PF 23-APR-1999; 99WO-CA00324.
XX
PR 23-APR-1998; 98US-0082791.
XX
PA (GENE-) GENESENSE TECHNOLOGIES INC.
XX
PI Wright JA, Young AH, Lee YS;
XX
DR WPI; 2000-023357/02.
XX
PT Antisense oligonucleotides that inhibit neuropilin expression, useful
PT for treating cancer -

```
XX
PS   Claim 4; Page 16; 57pp; English.
XX
CC   Sequences AAZ31431-460 represent antisense oligonucleotides which
CC   inhibit human neuropilin expression. The antisense oligonucleotides can
CC   be used to inhibit the growth or metastasis of a mammalian tumor and
CC   inhibit neovascularisation. The oligonucleotides may be used to treat
CC   various forms of cancers or tumors, such as sarcomas, melanomas,
CC   adenomas, carcinomas of solid tissue, hypoxic tumors, squamous cell
CC   carcinomas of the mouth, throat, larynx and lung, genitourinary cancers
CC   such as cervical and bladder cancer, hematopoietic cancers, colon cancer,
CC   breast cancer, pancreatic cancer, renal cancer, brain cancer, skin
CC   cancer, liver cancer, head and neck cancers, and nervous system cancers,
CC   as well as benign lesions such as papillomas. The methods may be used to
CC   treat neovascularisation disorders such as diabetic retinopathy, and
CC   retinopathy of prematurity and age related macular degeneration.
XX
SQ   Sequence 20 BP; 4 A; 7 C; 5 G; 4 T; 0 other;

  Query Match             100.0%;  Score 20;  DB 21;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 2.4;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CCCAGGGCACTCATGGCTAT 20
           ||||||||||||||||||||
Db       1 CCCAGGGCACTCATGGCTAT 20


RESULT 2
AAH62112/c
ID   AAH62112 standard; DNA; 21 BP.
XX
AC   AAH62112;
XX
DT   12-SEP-2001  (first entry)
XX
DE   Neuropilin 1 (NRP1) polymorphism containing DNA fragment #13.
XX
KW   Single nucleotide polymorphism; SNP; human; cancer; inflammation;
KW   heart disease; paternity testing; forensic science; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Variation       replace(11,T)
FT                   /*tag=  a
FT                   /standard_name= "single nucleotide polymorphism"
XX
PN   WO200138576-A2.
XX
PD   31-MAY-2001.
XX
PF   17-NOV-2000; 2000WO-US31639.
XX
PR   24-NOV-1999;   99US-0167334.
XX
PA   (WHED ) WHITEHEAD INST BIOMEDICAL RES.
XX
PI   Cargill M,  Ireland JS,  Lander ES;
XX
DR   WPI; 2001-367705/38.
XX
PT   New nucleic acid segments of the human genome, particularly from genes
PT   including polymorphic sites,for phenotype correlation, forensics,
PT   paternity testing, medicine and genetic analysis  -
XX
PS   Claim 1; Page 29; 80pp; English.
XX
CC   DNA sequences AAH62100 - AAH62688 represent segments of human genes which
CC   contain single nucleotide polymorphisms (SNPs). A method is included in
CC   the invention for analysing a nucleic acid sample, which consists of
CC   determining the base occupying any one of the polymorphic sites given in
CC   the SNP containing sequences. The nucleotide sequences can be used in the
CC   diagnosis or monitoring of diseases, such as cancer, inflammation, heart
CC   diseases, diseases of the cardiovascular system, and infection by
CC   microorganisms. The oligonucleotides are also useful in the manufacture
CC   of a medicament for the treatment or prophylaxis of the diseases, and as
CC   a pharmaceutical. SNP containing oligonucleotides are useful in
CC   applications such as phenotype correlation, forensics, paternity testing,
CC   medicine and genetic analysis.
XX
SQ   Sequence 21 BP; 3 A; 6 C; 9 G; 3 T; 0 other;

  Query Match             87.0%;  Score 17.4;  DB 22;  Length 21;
  Best Local Similarity   94.7%;  Pred. No. 44;
  Matches   18;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 CCCAGGGCACTCATGGCTA 19
           |||||||| ||||||||||
Db      19 CCCAGGGCGCTCATGGCTA 1


RESULT 3
AAZ31441
ID   AAZ31441 standard; DNA; 20 BP.
XX
AC   AAZ31441;
XX
DT   07-FEB-2000  (first entry)
XX
DE   Human neuropilin mRNA specific antisense oligo GTI3611.
XX
KW   Neuropilin; human; growth; metastasis; tumor; neovascularisation;
KW   cancer; papilloma; diabetic retinopathy; antisense; ss.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
PN   WO9955855-A2.
XX
PD   04-NOV-1999.
XX
PF   23-APR-1999;   99WO-CA00324.
XX
PR   23-APR-1998;   98US-0082791.
XX
PA   (GENE-) GENESENSE TECHNOLOGIES INC.
XX
PI   Wright JA,  Young AH,  Lee YS;
XX
DR   WPI; 2000-023357/02.
XX
PT   Antisense oligonucleotides that inhibit neuropilin expression, useful
PT   for treating cancer -
XX
PS   Claim 4; Page 16; 57pp; English.
XX
CC   Sequences AAZ31431-460 represent antisense oligonucleotides which
CC   inhibit human neuropilin expression. The antisense oligonucleotides can
CC   be used to inhibit the growth or metastasis of a mammalian tumor and
CC   inhibit neovascularisation. The oligonucleotides may be used to treat
CC   various forms of cancers or tumors, such as sarcomas, melanomas,
CC   adenomas, carcinomas of solid tissue, hypoxic tumors, squamous cell
CC   carcinomas of the mouth, throat, larynx and lung, genitourinary cancers
CC   such as cervical and bladder cancer, hematopoietic cancers, colon cancer,
CC   breast cancer, pancreatic cancer, renal cancer, brain cancer, skin
CC   cancer, liver cancer, head and neck cancers, and nervous system cancers,
CC   as well as benign lesions such as papillomas. The methods may be used to
CC   treat neovascularisation disorders such as diabetic retinopathy, and
CC   retinopathy of prematurity and age related macular degeneration.
XX
SQ   Sequence 20 BP; 4 A; 8 C; 6 G; 2 T; 0 other;

  Query Match             80.0%;  Score 16;  DB 21;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 2.1e+02;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy      1 CCCAGGGCACTCATGG 16
          ||||||||||||||||
Db      5 CCCAGGGCACTCATGG 20


RESULT 4
ABK88692/c
ID   ABK88692 standard; DNA; 39 BP.
XX
AC   ABK88692;
XX
DT   07-OCT-2002  (first entry)
XX
DE   Human transmembrane activator and CAML-interactor (TACI), PCR primer #2.
XX
KW   Human; tumour necrosis factor; TNF delta; pulmonary system disorder;
KW   immunoglobulin production; B-cell proliferation; immune system disorder;
KW   autoimmune disease; cancer; lymphoproliferative disorder; pain;
KW   microbial infection; parasitic infection; bone disease; atherosclerosis;
KW   cardiovascular disorder; neurodegenerative disease; wound healing;
KW   graft versus host disease; haematopoietic cell disorder; nephritis;
KW   transmembrane activator and CAML-interactor; TACI; TNF epsilon; PCR;
KW   primer; ss.
XX
OS   Homo sapiens.
XX
PN   US2002064829-A1.
XX
PD   30-MAY-2002.
XX
PF   14-JUN-2001; 2001US-0879919.
XX
PR   14-MAR-1996;   96US-016812P.
PR   15-JUN-2000; 2000US-211537P.
PR   23-OCT-2000; 2000US-241952P.
PR   13-DEC-2000; 2000US-254875P.
PR   16-MAR-2001; 2001US-276248P.
PR   23-MAR-2001; 2001US-277978P.
PR   25-MAY-2001; 2001US-293499P.
PR   12-MAR-1997;   97US-0815783.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Yu G,  Ni J,  Gentz RL,  Dillon PJ;
XX
DR   WPI; 2002-556722/59.
XX
PT   Novel human multimeric tumour necrosis factor delta or epsilon protein
PT   useful for treating disease or disorder of immune system such as
PT   autoimmune disease, immunodeficiency, or cancer of immune system  -
XX
PS   Example 29; Page 115; 143pp; English.
XX
CC   The present invention relates to the isolation of human tumour necrosis
CC   factor (TNF) delta and TNF epsilon proteins, and the polynucleotide
CC   sequences encoding them. The proteins are useful for modulating
CC   immunoglobulin production or for modulating proliferation of B-cells.
CC   The sequences of the invention are useful for treating diseases or
CC   disorders of the immune system. Such disorders include autoimmune
CC   diseases (e.g. systemic lupus erythematosus (SLE), acquired
CC   immunodeficiency syndrome (AIDS)), cancers of the immune system
CC   (e.g. chronic lymphocytic leukaemia (CLL), multiple myeloma,
CC   non-Hodgkin's lymphoma or Hodgkin's disease), lymphoproliferative
CC   disorders, microbial infections (e.g. viral, bacterial), parasitic
CC   infections, nephritis, bone disease (e.g. osteoporosis), atherosclerosis,
CC   pain, cardiovascular disorders (e.g. myocardial infarction, stroke),
CC   neurodegenerative diseases (e.g. Alzheimer's disease, Parkinson's
CC   disease), graft versus host disease, wound healing, haematopoietic cell
CC   disorders (e.g. anaemia), inflammatory disorders (e.g. asthma),
CC   diseases or disorders associated with various mucous membranes of the
CC   body (e.g. mucositis), and disorders of the pulmonary system. The
CC   proteins are also useful as a vaccine adjuvant that enhances immune
CC   responsiveness to specific antigens. The present sequence represents
CC   a PCR primer used to amplify human transmembrane activator and
CC   CAML-interactor (TACI) cDNA in the examples of the present invention.
XX
SQ   Sequence 39 BP; 8 A; 11 C; 13 G; 7 T; 0 other;

  Query Match           72.0%;  Score 14.4;  DB 24;  Length 39;
  Best Local Similarity   93.8%;  Pred. No. 1.3e+03;
  Matches   15;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 CCCAGGGCACTCATGG 16
          |||||| |||||||||
Db     32 CCCAGGCCACTCATGG 17


RESULT 5
ABK13410/c
ID   ABK13410 standard; DNA; 39 BP.
XX
AC   ABK13410;
XX
DT   09-APR-2002  (first entry)
XX
DE   Tumour necrosis factor (TNF) receptor TACI, 5' PCR primer #2.
XX
KW   Tumour necrosis factor; TNF; cytostatic; arteriosclerosis;
KW   analgesic; cerebroprotective; nootropic; neuroprotective; hepatotropic;
KW   immunoglobulin production; B cell proliferation; immunosuppressive;
KW   HIV; human immunodeficiency virus; autoimmune disease; immunodeficiency;
KW   Sjogren's syndrome; systemic lupus erythematosus; Hodgkin's disease;
KW   common variable immunodeficiency; CVID; non-Hodgkin's lymphoma; AIDS;
KW   acquired immunodeficiency virus; cancer; multiple myeloma; CLL;
KW   chronic lymphocytic leukaemia; lymphoproliferative disorder;
KW   bacterial infection; viral infection; osteoporosis; atherosclerosis;
KW   pain; cardiovascular disorder; stroke; allergy; Alzheimer's disease;
KW   neurodegenerative disease; inflammation; liver disease; cirrhosis;
KW   cardiomyopathy; diabetes; asthma; psoriasis; glomerulonephritis;
KW   ulcerative colitis; angiogenesis; septic shock; wound healing;
KW   PCR; primer; ss; tumour necrosis factor receptor; TACI.
XX
OS   Homo sapiens.
XX
PN   WO200196528-A2.
XX
PD   20-DEC-2001.
XX
PF   14-JUN-2001; 2001WO-US19026.
XX
PR   15-JUN-2000; 2000US-211537P.
PR   23-OCT-2000; 2000US-241952P.
PR   13-DEC-2000; 2000US-254875P.
PR   16-MAR-2001; 2001US-276248P.
PR   23-MAR-2001; 2001US-277978P.
PR   25-MAY-2001; 2001US-293499P.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Yu G,  Ni J,  Gentz RL,  Dillon PJ,  Hilbert D;
XX
DR   WPI; 2002-130727/17.
XX
PT   Novel multimeric human tumour necrosis factor delta or epsilon protein
PT   useful for treating cancer, immune system disorders, infection,
PT   cardiovascular disorders, liver disease, cardiomyopathy, diabetes and
PT   psoriasis  -
XX
PS   Example 29; Page 295; 344pp; English.
XX
CC   The invention describes a multimeric human tumour necrosis factor (TNF)
CC   delta or epsilon protein (I). (I) or a composition containing them (II)
CC   are useful for modulating immunoglobulin production or proliferation of B
CC   cells. (I) or (II) is useful: for treating a disease or disorder of the
CC   immune system, preferably an autoimmune disease (e.g. Sjogren's syndrome,
```

```
CC   systemic lupus erythematosus or common variable immunodeficiency (CVID));
CC   an immunodeficiency e.g. acquired immunodeficiency syndrome (AIDS);
CC   cancer of the immune system (e.g. Hodgkin's disease, non-Hodgkin's
CC   lymphoma, multiple myeloma and chronic lymphocytic leukaemia (CLL)); in
CC   the diagnosis and treatment or prevention of cancer, lymphoproliferative
CC   disorder, bacterial and viral infections, osteoporosis, atherosclerosis,
CC   pain, cardiovascular disorders (e.g. stroke), allergy, inflammation,
CC   neurodegenerative disease (e.g. Alzheimer's disease), liver disease (e.g.
CC   cirrhosis), cardiomyopathy, diabetes, asthma, psoriasis, septic shock,
CC   glomerulonephritis, ulcerative colitis, arteriosclerosis; for promoting
CC   angiogenesis and wound healing; as a diagnostic research reagent; as an
CC   agent to target and kill cells expressing a TNFdelta and/or TNFepsilon
CC   receptor; in apoptosis of transformed cell lines; mediation of cell
CC   activation and proliferation; and as an immunogen to produce (II). (II)
CC   is useful to purify, detect and target (I), for measuring levels of (I)
CC   in biological samples, for immunophenotyping samples, and to treat,
CC   inhibit or prevent diseases and disorders associated with aberrant
CC   expression and/or activity of (I). This sequence represents a 5' PCR
CC   primer used to isolate the TNF receptor superfamily member TACI, required
CC   in an assay to establish if tumour necrosis factor epsilon (TNF-epsilon)
CC   binds the receptor described in the method of the invention.
XX
SQ   Sequence 39 BP; 8 A; 11 C; 13 G; 7 T; 0 other;

  Query Match             72.0%;  Score 14.4;  DB 24;  Length 39;
  Best Local Similarity   93.8%;  Pred. No. 1.3e+03;
  Matches   15;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 CCCAGGGCACTCATGG 16
           |||||| |||||||||
Db      32 CCCAGGCCACTCATGG 17


RESULT 6
AAT43862
ID   AAT43862 standard; DNA; 23 BP.
XX
AC   AAT43862;
XX
DT   21-MAR-1997  (first entry)
XX
DE   Mouse ob gene reverse PCR primer 559.
XX
KW   Obesity; ob protein; ob receptor; gene therapy; diabetes;
KW   polymerase chain reaction; PCR; primer; ss.
XX
OS   Synthetic.
XX
PN   WO9635787-A1.
XX
PD   14-NOV-1996.
XX
PF   08-MAY-1996;   96WO-US06609.
XX
PR   08-MAY-1995;   95US-0437834.
XX
PA   (CHIR ) CHIRON CORP.
XX
PI   Giese KW,  Williams LT;
XX
DR   WPI; 1996-518675/51.
XX
PT   Nucleic acid encoding obesity protein - useful for treating obesity
PT   or associated diseases e.g. type II diabetes.
XX
PS   Example 2; Page 38; 66pp; English.
XX
CC   PCR reverse primer 559 (AAT43862) was used with forward primer 560
CC   (AAT43861) to amplify DNA encoding amino acids 22-167 (AAW07505) of
CC   the mouse obesity (ob) protein.  Full-length ob cDNA was used as
CC   template.  For expression in prokaryotes, the amplified DNA fragment
CC   was ligated into vector pHB40P contg. the sequence for heart muscle
CC   kinase and the Myc epitope, to give DNA construct #1150 (see also
CC   AAT43856).
XX
SQ   Sequence 23 BP; 4 A; 8 C; 8 G; 3 T; 0 other;

  Query Match             71.0%;  Score 14.2;  DB 17;  Length 23;
  Best Local Similarity   84.2%;  Pred. No. 1.5e+03;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 CCCAGGGCACTCATGGCTA 19
           ||| ||||| ||| |||||
Db       3 CCCGGGGCATTCAGGGCTA 21


RESULT 7
AAL27305
ID   AAL27305 standard; DNA; 51 BP.
XX
AC   AAL27305;
XX
DT   24-JAN-2002  (first entry)
XX
DE   Human SNP oligonucleotide #513.
XX
KW   Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
KW   neuroprotective; antimicrobial; gene therapy; vaccine; amylase; cancer;
KW   amyloid protein; angiopoietin; apoptosis related protein; cadherin;
KW   cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
KW   complement related protein; cytochrome; kinesin; cytokine; interferon;
KW   interleukin; G-protein coupled receptor; thioesterase; inflammation;
KW   multifactorial disease; autoimmune disease; infection;
KW   nervous system disease; ss.
XX
OS   Homo sapiens.
XX
PN   WO200147944-A2.
XX
PD   05-JUL-2001.
XX
PF   28-DEC-2000; 2000WO-US35498.
XX
PR   28-DEC-1999;   99US-0173419.
PR   27-DEC-2000; 2000US-0173419.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Shimkets RA,  Leach M;
XX
DR   WPI; 2001-465210/50.
XX
PT   Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
PT   oncogenes and histones, useful for diagnosing and treating, e.g.
PT   cancer, autoimmune diseases and infections -
XX
PS   Claim 1; Page 1536; 4143pp; English.
XX
CC   The present invention relates to oligonucleotides encoding polymorphic
CC   variants of proteins related to amylases, amyloid proteins, angiopoietin,
CC   apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,
CC   histones, kinases, colony stimulating factors, complement related
CC   proteins, cytochromes, kinesins, cytokines, interferons, interleukins,
CC   G-protein coupled receptors and thioesterases. The present sequence is
CC   one such oligonucleotide. The oligonucleotides and the peptides encoded
CC   by them may be used in the prevention, diagnosis and treatment of
CC   diseases associated with inappropriate expression of the proteins listed
CC   above. Disorders that may be prevented, diagnosed and/or treated include
CC   multifactorial diseases with a genetic component, such as autoimmune
CC   diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,
CC   systemic lupus erythromatosus and Grave's disease), inflammation, cancer
CC   (e.g. cancers of the bladder, brain, breast, colon and kidney,
CC   leukaemia), diseases of the nervous system and an infection of pathogenic
CC   organisms.
XX
SQ   Sequence 51 BP; 12 A; 16 C; 11 G; 12 T; 0 other;
```

```
  Query Match             71.0%;  Score 14.2;  DB 22;  Length 51;
  Best Local Similarity   84.2%;  Pred. No. 1.6e+03;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 CCCAGGGCACTCATGGCTA 19
          | ||||||| |||||| ||
Db     33 CTCAGGGCAGTCATGGATA 51


RESULT 8
AAL27306
ID   AAL27306 standard; DNA; 51 BP.
XX
AC   AAL27306;
XX
DT   24-JAN-2002  (first entry)
XX
DE   Human SNP oligonucleotide #514.
XX
KW   Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
KW   neuroprotective; antimicrobial; gene therapy; vaccine; amylase; cancer;
KW   amyloid protein; angiopoietin; apoptosis related protein; cadherin;
KW   cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
KW   complement related protein; cytochrome; kinesin; cytokine; interferon;
KW   interleukin; G-protein coupled receptor; thioesterase; inflammation;
KW   multifactorial disease; autoimmune disease; infection;
KW   nervous system disease; ss.
XX
OS   Homo sapiens.
XX
PN   WO200147944-A2.
XX
PD   05-JUL-2001.
XX
PF   28-DEC-2000; 2000WO-US35498.
XX
PR   28-DEC-1999;   99US-0173419.
PR   27-DEC-2000; 2000US-0173419.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Shimkets RA,  Leach M;
XX
DR   WPI; 2001-465210/50.
XX
PT   Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
PT   oncogenes and histones, useful for diagnosing and treating, e.g.
PT   cancer, autoimmune diseases and infections -
XX
PS   Claim 1; Page 1536; 4143pp; English.
XX
CC   The present invention relates to oligonucleotides encoding polymorphic
CC   variants of proteins related to amylases, amyloid proteins, angiopoietin,
CC   apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,
CC   histones, kinases, colony stimulating factors, complement related
CC   proteins, cytochromes, kinesins, cytokines, interferons, interleukins,
CC   G-protein coupled receptors and thioesterases. The present sequence is
CC   one such oligonucleotide. The oligonucleotides and the peptides encoded
CC   by them may be used in the prevention, diagnosis and treatment of
CC   diseases associated with inappropriate expression of the proteins listed
CC   above. Disorders that may be prevented, diagnosed and/or treated include
CC   multifactorial diseases with a genetic component, such as autoimmune
CC   diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,
CC   systemic lupus erythromatosus and Grave's disease), inflammation, cancer
CC   (e.g. cancers of the bladder, brain, breast, colon and kidney,
CC   leukaemia), diseases of the nervous system and an infection of pathogenic
CC   organisms.
XX
SQ   Sequence 51 BP; 10 A; 18 C; 11 G; 12 T; 0 other;

  Query Match             71.0%;  Score 14.2;  DB 22;  Length 51;
  Best Local Similarity   84.2%;  Pred. No. 1.6e+03;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 CCCAGGGCACTCATGGCTA 19
          | ||||||| |||||| ||
Db     30 CTCAGGGCAGTCATGGATA 48


RESULT 9
AAL27308
ID   AAL27308 standard; DNA; 51 BP.
XX
AC   AAL27308;
XX
DT   24-JAN-2002  (first entry)
XX
DE   Human SNP oligonucleotide #516.
XX
KW   Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
KW   neuroprotective; antimicrobial; gene therapy; vaccine; amylase; cancer;
KW   amyloid protein; angiopoietin; apoptosis related protein; cadherin;
KW   cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
KW   complement related protein; cytochrome; kinesin; cytokine; interferon;
KW   interleukin; G-protein coupled receptor; thioesterase; inflammation;
KW   multifactorial disease; autoimmune disease; infection;
KW   nervous system disease; ss.
XX
OS   Homo sapiens.
XX
PN   WO200147944-A2.
XX
PD   05-JUL-2001.
XX
PF   28-DEC-2000; 2000WO-US35498.
XX
PR   28-DEC-1999;   99US-0173419.
PR   27-DEC-2000; 2000US-0173419.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Shimkets RA,  Leach M;
XX
DR   WPI; 2001-465210/50.
XX
PT   Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
PT   oncogenes and histones, useful for diagnosing and treating, e.g.
PT   cancer, autoimmune diseases and infections -
XX
PS   Claim 1; Page 1537; 4143pp; English.
XX
CC   The present invention relates to oligonucleotides encoding polymorphic
CC   variants of proteins related to amylases, amyloid proteins, angiopoietin,
CC   apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,
CC   histones, kinases, colony stimulating factors, complement related
CC   proteins, cytochromes, kinesins, cytokines, interferons, interleukins,
CC   G-protein coupled receptors and thioesterases. The present sequence is
CC   one such oligonucleotide. The oligonucleotides and the peptides encoded
CC   by them may be used in the prevention, diagnosis and treatment of
CC   diseases associated with inappropriate expression of the proteins listed
CC   above. Disorders that may be prevented, diagnosed and/or treated include
CC   multifactorial diseases with a genetic component, such as autoimmune
CC   diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,
CC   systemic lupus erythromatosus and Grave's disease), inflammation, cancer
CC   (e.g. cancers of the bladder, brain, breast, colon and kidney,
CC   leukaemia), diseases of the nervous system and an infection of pathogenic
CC   organisms.
XX
SQ   Sequence 51 BP; 10 A; 16 C; 11 G; 14 T; 0 other;

  Query Match             71.0%;  Score 14.2;  DB 22;  Length 51;
  Best Local Similarity   84.2%;  Pred. No. 1.6e+03;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 CCCAGGGCACTCATGGCTA 19
```

```
          | |||||||| |||||| ||
Db      3 CTCAGGGCAGTCATGGATA 21


RESULT 10
ABN39029/c
ID   ABN39029 standard; DNA; 60 BP.
XX
AC   ABN39029;
XX
DT   15-JUL-2002  (first entry)
XX
DE   Human spliced transcript detection oligonucleotide SEQ ID NO:11777.
XX
KW   Human; mouse; rat; splice transcript; detection; RNA transcript;
KW   splice variant; transcriptome; oligonucleotide library; ss.
XX
OS   Homo sapiens.
XX
PN   WO200210449-A2.
XX
PD   07-FEB-2002.
XX
PF   20-JUL-2001; 2001WO-IB01903.
XX
PR   28-JUL-2000; 2000US-221607P.
PR   02-MAY-2001; 2001US-287724P.
XX
PA   (COMP-) COMPUGEN INC.
XX
PI   Shoshan A,  Wasserman A,  Mintz E,  Mintz L,  Faigler S;
XX
DR   WPI; 2002-257383/30.
XX
PT   New oligonucleotide libraries comprising oligonucleotides which
PT   selectively hybridize to mRNAs transcribed from a transcription unit of
PT   a genome, useful for detecting tissue-, pathology-, and
PT   developmental-specific genes  -
XX
PS   Example 1; SEQ ID 11777; 47pp; English.
XX
CC   The present invention describes oligonucleotide libraries for detecting
CC   messenger RNAs that populate a (sub-)transcriptome, where the
CC   (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC   transcription units that populate a genome. The library comprises
CC   several oligonucleotides, each capable of hybridising selectively to a
CC   set of messenger RNAs transcribed from a given transcription unit of
CC   the genome, which encodes one or more messenger RNA splice variants.
CC   The oligonucleotide libraries are useful for detecting mRNAs from a
CC   biological sample, in expression profiling studies, in qualitatively or
CC   quantitatively characterising the corresponding transcriptome, and in
CC   detecting RNA transcripts and splice variants of human or animal
CC   transcriptomes. The libraries may also be used as specialised mini
CC   libraries to detect transcripts of a sub-transcriptome under a
CC   particular biological or pathological state, and so allowing the
CC   detection of tissue- and pathology-specific genes such as those genes
CC   only expressed in specific tissue under a specific pathological
CC   condition; to detect developmental specific genes; and to detect RNA
CC   transcripts and splice variants of a transcriptome of a patient suffering
CC   from a particular disorder. ABN27253 to ABN59589 represent
CC   oligonucleotide sequences from rats, humans and mice, which are used in
CC   the exemplification of the present invention.
CC   N.B. The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 60 BP; 13 A; 13 C; 17 G; 17 T; 0 other;

  Query Match             71.0%;  Score 14.2;  DB 24;  Length 60;
  Best Local Similarity   84.2%;  Pred. No. 1.7e+03;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 CCCAGGGCACTCATGGCTA 19
```

```
            |||| ||||  ||||||||
Db      22 CCCATGGCAGACATGGCTA 4


RESULT 11
AAF74885/c
ID   AAF74885 standard; DNA; 21 BP.
XX
AC   AAF74885;
XX
DT   22-MAY-2001  (first entry)
XX
DE   Chicken neuropilin-1 mutagenesis PCR primer SEQ ID NO:8.
XX
KW   Chicken; neuropilin-1; secreted semaphorin receptor; collapse; motility;
KW   axon growth core; axon generation; mutagenesis; PCR primer; ss.
XX
OS   Gallus gallus.
XX
PN   WO200118173-A2.
XX
PD   15-MAR-2001.
XX
PF   08-SEP-2000; 2000WO-US24635.
XX
PR   10-SEP-1999;   99US-0153309.
PR   16-DEC-1999;   99US-0171176.
XX
PA   (UYPE-) UNIV PENNSYLVANIA.
XX
PI   Raper JA,  Renzi MJ;
XX
DR   WPI; 2001-235194/24.
XX
PT   New dominant negative neuropilin-1 receptor, useful for modulating or
PT   inactivating activity of selected secreted semaphorins and inhibiting
PT   or preventing collapse or motility of axon growth cone  -
XX
PS   Example 1; Page 28; 68pp; English.
XX
CC   The present invention describes an isolated DNA encoding a dominant
CC   negative receptor (I), where the DNA comprises a nucleic acid sequence
CC   encoding a neuropilin-1 which has semaphorin receptor specific
CC   antigenicity or immunogenicity, including homologues, modifications,
CC   derivatives and active fragments. Also described is a protein comprising
CC   a dominant negative receptor, which has receptor-specific antigenicity
CC   or immunogenicity for semaphorin 3A and semaphorin 3C, but not for
CC   semaphorin 3F. (I) Is useful for modulating or inactivating the activity
CC   of selected secreted semaphorins, for inhibiting or preventing the
CC   collapse or motility of an axon growth cone, where motility or collapse
CC   is mediated by a secreted semaphorin, for modulating overgrowth or
CC   premature entry of axons to their targets in vivo, and for enhancing
CC   axon generation or regeneration by blocking secreted semaphorin
CC   binding, by adding or overexpressing (I). The axon growth occurs in a
CC   developing or regenerating neurological system. The present sequence
CC   represents a PCR primer which is used in an example from the present
CC   invention.
XX
SQ   Sequence 21 BP; 8 A; 5 C; 4 G; 4 T; 0 other;

  Query Match             70.0%;  Score 14;  DB 22;  Length 21;
  Best Local Similarity   100.0%;  Pred. No. 1.9e+03;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       7 GCACTCATGGCTAT 20
           ||||||||||||||
Db      20 GCACTCATGGCTAT 7


RESULT 12
AAF74884
ID   AAF74884 standard; DNA; 42 BP.
XX
```

```
AC   AAF74884;
XX
DT   22-MAY-2001  (first entry)
XX
DE   Chicken neuropilin-1 mutagenesis PCR primer SEQ ID NO:7.
XX
KW   Chicken; neuropilin-1; secreted semaphorin receptor; collapse; motility;
KW   axon growth core; axon generation; mutagenesis; PCR primer; ss.
XX
OS   Gallus gallus.
XX
PN   WO200118173-A2.
XX
PD   15-MAR-2001.
XX
PF   08-SEP-2000; 2000WO-US24635.
XX
PR   10-SEP-1999;   99US-0153309.
PR   16-DEC-1999;   99US-0171176.
XX
PA   (UYPE-) UNIV PENNSYLVANIA.
XX
PI   Raper JA,  Renzi MJ;
XX
DR   WPI; 2001-235194/24.
XX
PT   New dominant negative neuropilin-1 receptor, useful for modulating or
PT   inactivating activity of selected secreted semaphorins and inhibiting
PT   or preventing collapse or motility of axon growth cone  -
XX
PS   Example 1; Page 28; 68pp; English.
XX
CC   The present invention describes an isolated DNA encoding a dominant
CC   negative receptor (I), where the DNA comprises a nucleic acid sequence
CC   encoding a neuropilin-1 which has semaphorin receptor specific
CC   antigenicity or immunogenicity, including homologues, modifications,
CC   derivatives and active fragments. Also described is a protein comprising
CC   a dominant negative receptor, which has receptor-specific antigenicity
CC   or immunogenicity for semaphorin 3A and semaphorin 3C, but not for
CC   semaphorin 3F. (I) Is useful for modulating or inactivating the activity
CC   of selected secreted semaphorins, for inhibiting or preventing the
CC   collapse or motility of an axon growth cone, where motility or collapse
CC   is mediated by a secreted semaphorin, for modulating overgrowth or
CC   premature entry of axons to their targets in vivo, and for enhancing
CC   axon generation or regeneration by blocking secreted semaphorin
CC   binding, by adding or overexpressing (I). The axon growth occurs in a
CC   developing or regenerating neurological system. The present sequence
CC   represents a PCR primer which is used in an example from the present
CC   invention.
XX
SQ   Sequence 42 BP; 9 A; 9 C; 12 G; 12 T; 0 other;

  Query Match             70.0%;  Score 14;  DB 22;  Length 42;
  Best Local Similarity   100.0%;  Pred. No. 2e+03;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       7 GCACTCATGGCTAT 20
           ||||||||||||||
Db       2 GCACTCATGGCTAT 15


RESULT 13
ABQ61256
ID   ABQ61256 standard; DNA; 20 BP.
XX
AC   ABQ61256;
XX
DT   03-OCT-2002  (first entry)
XX
DE   Human aquaporin 5 (AQP5) exon 1 PCR primer 4.
XX
KW   Human; ss; PCR; primer; aquaporin; AQP5; AQP; water channel protein;
KW   oligonucleotide chip; OGN chip; cDNA chip; lung cancer;
KW   mutation detection; polymorphism detection; gene expression.
XX
OS   Homo sapiens.
XX
PN   WO200220787-A1.
XX
PD   14-MAR-2002.
XX
PF   10-SEP-2001; 2001WO-KR01528.
XX
PR   09-SEP-2000; 2000KR-0053821.
XX
PA   (GOOD-) GOODGENE INC.
PA   (MOON/) MOON W.
PA   (MOON/) MOON C.
XX
PI   Moon W,  Moon C,  Moon Y,  Kim B,  Kim D,  Shin C,  Um T,  Kim H;
PI   Song M,  Kim H,  Song S;
XX
DR   WPI; 2002-393847/42.
XX
PT   Novel aquaporin 5 gene mutant useful for diagnosing lung, stomach,
PT   colon, prostate, or head or neck cancer  -
XX
PS   Example 9; Page 151; 154pp; English.
XX
CC   The invention comprises a mutant form of the human aquaporin 5 (AQP5)
CC   gene. Aquaporin (AQP) is a family of water channel proteins, through
CC   which water is transported into and out of cells - ten types of mammalian
CC   AQP have been identified so far. The invention also comprises an
CC   oligonucleotide (OGN) chip having 902 oligonucleotide primer sequences
CC   and a cDNA chip comprising one or more sequences from the human AQP5
CC   gene. The mutant AQP5 gene is useful for diagnosing cancer (i.e lung
CC   cancer). The OGN chip is useful for detecting mutations and polymorphisms
CC   in AQP5 and the cDNA chip is useful for analysis of gene expression. The
CC   present DNA sequence represents a human aquaporin (AQP) gene PCR primer.
XX
SQ   Sequence 20 BP; 3 A; 9 C; 5 G; 3 T; 0 other;

  Query Match             69.0%;  Score 13.8;  DB 24;  Length 20;
  Best Local Similarity   88.2%;  Pred. No. 2.4e+03;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 CCCAGGGCACTCATGGC 17
           |||||||||||||  ||
Db       1 CCCAGGGCACTCACCGC 17


RESULT 14
ABI84610/c
ID   ABI84610 standard; DNA; 24 BP.
XX
AC   ABI84610;
XX
DT   15-FEB-2002  (first entry)
XX
DE   Capture oligonucleotide Zip ID#1107 oligo #1.
XX
KW   Human; K-ras; PCR primer; probe; capture probe; mutation detection;
KW   ligase detection reaction; LDR; p53; BRCA1; BRCA2; infectious disease;
KW   infection; 21 hydroxylase deficiency; Turner Syndrome; obesity;
KW   cancer; oncogene; tumour suppressor; human papillomavirus; forensic;
KW   environmental monitoring; food industry; feed industry; ss.
XX
OS   Synthetic.
XX
PN   WO200179548-A2.
XX
PD   25-OCT-2001.
XX
PF   04-APR-2001; 2001WO-US10958.
XX
PR   14-APR-2000; 2000US-197271P.
```

```
XX
PA   (CORR ) CORNELL RES FOUND INC.
XX
PI   Barany F,  Zirvi M,  Gerry NP,  Favis R,  Kliman R;
XX
DR   WPI; 2002-034366/04.
XX
PT   Designing capture oligonucleotide probes for use on a support to which
PT   complementary oligonucleotides hybridize with little mismatch -
XX
PS   Example 5; Fig 25; 300pp; English.
XX
CC   The present invention describes a method (M1) for designing capture
CC   oligonucleotide probes (I) for use on a support to which complementary
CC   oligonucleotide probes (II) will hybridise with little mismatch, where
CC   (I) have melting temperatures within a narrow range. The method is useful
CC   for detecting infectious diseases caused by bacterial infectious agents
CC   e.g. Salmonella, Listeria monocytogenes and Haemophilus influenza, fungal
CC   infectious agents e.g. Cryptococcus neoformans, Candida albicans and
CC   Aspergillus fumigautus, viruses e.g. T-cell lymphocytotrophis cirus,
CC   Epstein-Barr virus and polio virus, and parasitic infectious agents
CC   selected from Onchoverva volvulus, Entamoeba histolytica and Dracunculus
CC   medinesis. The method is also useful for detecting genetic diseases such
CC   as 21 hydroxylase deficiency, Turner Syndrome and obesity defects.
CC   Detecting cancer involving oncogenes, tumour suppressor genes, or genes
CC   involved in DNA amplification, replication, recombination or repair, the
CC   cancer is specifically associated with a gene selected from BRCA1 gene,
CC   p53 gene, human papillomavirus types 16 and 18 and liver cancers. The
CC   method is also used for environmental monitoring, forensics and the food
CC   and feed industry, detecting comprises scanning (using e.g. a scanning
CC   electron microscope and infrared microscope) the support at the
CC   particular sites and identifying if ligation of the oligonucleotide probe
CC   sets occurred and correlating (using a computer) identified ligation to a
CC   presence or absence of the target nucleotide sequences. ABI82074 to
CC   ABI97546 represent oligonucleotide sequences used in the exemplification
CC   of the present invention.
XX
SQ   Sequence 24 BP; 3 A; 7 C; 7 G; 7 T; 0 other;

  Query Match             69.0%;  Score 13.8;  DB 24;  Length 24;
  Best Local Similarity   88.2%;  Pred. No. 2.4e+03;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       3 CAGGGCACTCATGGCTA 19
           || |||||||| |||||
Db      20 CAAGGCACTCAAGGCTA 4


RESULT 15
ABI84611
ID   ABI84611 standard; DNA; 24 BP.
XX
AC   ABI84611;
XX
DT   15-FEB-2002  (first entry)
XX
DE   Capture oligonucleotide Zip ID#1107 oligo #2.
XX
KW   Human; K-ras; PCR primer; probe; capture probe; mutation detection;
KW   ligase detection reaction; LDR; p53; BRCA1; BRCA2; infectious disease;
KW   infection; 21 hydroxylase deficiency; Turner Syndrome; obesity;
KW   cancer; oncogene; tumour suppressor; human papillomavirus; forensic;
KW   environmental monitoring; food industry; feed industry; ss.
XX
OS   Synthetic.
XX
PN   WO200179548-A2.
XX
PD   25-OCT-2001.
XX
PF   04-APR-2001; 2001WO-US10958.
XX
PR   14-APR-2000; 2000US-197271P.
```


```
XX
PA   (CORR ) CORNELL RES FOUND INC.
XX
PI   Barany F,  Zirvi M,  Gerry NP,  Favis R,  Kliman R;
XX
DR   WPI; 2002-034366/04.
XX
PT   Designing capture oligonucleotide probes for use on a support to which
PT   complementary oligonucleotides hybridize with little mismatch -
XX
PS   Example 5; Fig 25; 300pp; English.
XX
CC   The present invention describes a method (M1) for designing capture
CC   oligonucleotide probes (I) for use on a support to which complementary
CC   oligonucleotide probes (II) will hybridise with little mismatch, where
CC   (I) have melting temperatures within a narrow range. The method is useful
CC   for detecting infectious diseases caused by bacterial infectious agents
CC   e.g. Salmonella, Listeria monocytogenes and Haemophilus influenza, fungal
CC   infectious agents e.g. Cryptococcus neoformans, Candida albicans and
CC   Aspergillus fumigautus, viruses e.g. T-cell lymphocytotrophis cirus,
CC   Epstein-Barr virus and polio virus, and parasitic infectious agents
CC   selected from Onchoverva volvulus, Entamoeba histolytica and Dracunculus
CC   medinesis. The method is also useful for detecting genetic diseases such
CC   as 21 hydroxylase deficiency, Turner Syndrome and obesity defects.
CC   Detecting cancer involving oncogenes, tumour suppressor genes, or genes
CC   involved in DNA amplification, replication, recombination or repair, the
CC   cancer is specifically associated with a gene selected from BRCA1 gene,
CC   p53 gene, human papillomavirus types 16 and 18 and liver cancers. The
CC   method is also used for environmental monitoring, forensics and the food
CC   and feed industry, detecting comprises scanning (using e.g. a scanning
CC   electron microscope and infrared microscope) the support at the
CC   particular sites and identifying if ligation of the oligonucleotide probe
CC   sets occurred and correlating (using a computer) identified ligation to a
CC   presence or absence of the target nucleotide sequences. ABI82074 to
CC   ABI97546 represent oligonucleotide sequences used in the exemplification
CC   of the present invention.
XX
```

```
SQ   Sequence 24 BP; 7 A; 7 C; 7 G; 3 T; 0 other;

  Query Match             69.0%;  Score 13.8;  DB 24;  Length 24;
  Best Local Similarity   88.2%;  Pred. No. 2.4e+03;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       3 CAGGGCACTCATGGCTA 19
           || |||||||| |||||
Db       5 CAAGGCACTCAAGGCTA 21


Search completed: November 23, 2002, 06:29:32
Job time : 100.6 secs
```

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 05:53:51 ; Search time 21.55 Seconds
(without alignments)
284.619 Million cell updates/sec

Title: US-09-296-264-20
Perfect score: 20
Sequence: 1 cccagggcactcatggctat 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 687286

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	13.2	66.0	24	2	US-08-288-508C-10	Sequence 10, Appl
c	2	13.2	66.0	24	3	US-09-005-165-8	Sequence 8, Appli
	3	13.2	66.0	24	3	US-08-289-222E-20	Sequence 20, Appl
	4	13.2	66.0	24	4	US-09-054-526B-20	Sequence 20, Appl
	5	12.8	64.0	57	1	US-08-242-663A-5	Sequence 5, Appli
	6	12.8	64.0	57	2	US-08-573-890-3	Sequence 3, Appli
	7	12.8	64.0	57	4	US-09-248-179-3	Sequence 3, Appli
	8	12.8	64.0	57	5	PCT-US95-06132-5	Sequence 5, Appli
c	9	12.8	64.0	66	3	US-07-728-220C-14	Sequence 14, Appl
c	10	12.8	64.0	69	1	US-08-352-179-10	Sequence 10, Appl
	11	12.6	63.0	23	2	US-08-823-516-105	Sequence 105, App
	12	12.6	63.0	23	3	US-08-759-038-90	Sequence 90, Appl
	13	12.6	63.0	23	3	US-08-758-314-90	Sequence 90, Appl
	14	12.6	63.0	23	3	US-08-611-587-32	Sequence 32, Appl
c	15	12.6	63.0	29	3	US-08-526-136-5	Sequence 5, Appli
c	16	12.6	63.0	61	4	US-09-513-783A-145	Sequence 145, App
c	17	12.6	63.0	69	4	US-08-483-511-77	Sequence 77, Appl
c	18	12.6	63.0	96	4	US-08-952-793-276	Sequence 276, App
c	19	12.6	63.0	96	5	PCT-US96-09455A-276	Sequence 276, App
	20	12.4	62.0	18	1	US-08-164-200-14	Sequence 14, Appl
c	21	12.4	62.0	18	3	US-09-040-025-39	Sequence 39, Appl
c	22	12.4	62.0	18	4	US-09-040-025-39	Sequence 39, Appl
c	23	12.4	62.0	37	4	US-09-263-904-7	Sequence 7, Appli
	24	12.4	62.0	38	2	US-08-292-620A-2308	Sequence 2308, Ap
	25	12.4	62.0	38	3	US-09-071-845-2308	Sequence 2308, Ap
c	26	12.2	61.0	22	1	US-08-479-723A-87	Sequence 87, Appl
c	27	12.2	61.0	23	1	US-08-479-723A-78	Sequence 78, Appl
	28	12.2	61.0	23	4	US-09-068-805-8	Sequence 8, Appli
c	29	12.2	61.0	25	4	US-09-182-728A-4	Sequence 4, Appli
c	30	12.2	61.0	25	4	US-09-795-232-4	Sequence 4, Appli
	31	12.2	61.0	28	4	US-09-301-374-8	Sequence 8, Appli
c	32	12.2	61.0	30	4	US-08-809-326A-23	Sequence 23, Appl
c	33	12.2	61.0	30	4	US-08-891-292A-60	Sequence 60, Appl
	34	12.2	61.0	31	3	US-08-946-914-55	Sequence 55, Appl
	35	12.2	61.0	31	4	US-09-656-450-55	Sequence 55, Appl
	36	12.2	61.0	32	3	US-08-946-914-57	Sequence 57, Appl
c	37	12.2	61.0	32	4	US-08-891-292A-64	Sequence 64, Appl
	38	12.2	61.0	32	4	US-09-656-450-57	Sequence 57, Appl
c	39	12.2	61.0	35	4	US-08-809-326A-11	Sequence 11, Appl
	40	12.2	61.0	35	4	US-08-809-326A-12	Sequence 12, Appl
c	41	12.2	61.0	49	1	US-07-753-110B-10	Sequence 10, Appl
c	42	12.2	61.0	49	1	US-08-503-730-4	Sequence 4, Appli
c	43	12.2	61.0	49	2	US-08-507-634-11	Sequence 11, Appl
	44	12.2	61.0	55	1	US-07-972-032-76	Sequence 76, Appl
c	45	12.2	61.0	56	2	US-08-823-516-103	Sequence 103, App

ALIGNMENTS

```
RESULT 1
US-08-288-508C-10
; Sequence 10, Application US/08288508C
; Patent No. 5994094
;  GENERAL INFORMATION:
;    APPLICANT:  H tten, Gertrud
;    APPLICANT:  Neidhardt, Helge
;    APPLICANT:  Paulista, Michael
;    TITLE OF INVENTION:  NEW GROWTH/DIFFERENTIATING FACTOR OF
;    TITLE OF INVENTION:  THE TGF-  FAMILY
;    NUMBER OF SEQUENCES:  40
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Nikaido, Marmelstein, Murray & Oram LLP
;      STREET:  655 Fifteenth Street N.W. Suite 330
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  U.S.A.
;      ZIP:  20005-5701
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/288,508C
;      FILING DATE:  10-AUG-1994
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  DE P 43 26 829.3
;      FILING DATE:  10-AUG-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  DE P 44 18 222.8
;      FILING DATE:  25-MAY-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  DE P 44 20 157.5
;      FILING DATE:  09-JUN-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  JAHNS, Kristina M.
;      REGISTRATION NUMBER:  P-41,092
;      REFERENCE/DOCKET NUMBER:  P564-4019
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)638-5000
;      TELEFAX:  (202)638-4810
;  INFORMATION FOR SEQ ID NO:  10:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  24 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA
```

```
US-08-288-508C-10

  Query Match             66.0%;  Score 13.2;  DB 2;  Length 24;
  Best Local Similarity   83.3%;  Pred. No. 5.9e+02;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 CCAGGGCACTCATGGCTA 19
           |||||||||| ||| | |
Db       2 CCAGGGCACTAATGTCAA 19


RESULT 2
US-09-005-165-8/c
; Sequence 8, Application US/09005165
; Patent No. 6093542
;  GENERAL INFORMATION:
;    APPLICANT:  ROMANO, JOSEPH
;    APPLICANT:  SHURTLIFF, ROXANNE
;    APPLICANT:  PAL, RANAJIT
;    TITLE OF INVENTION:  ISOTHERMAL TRANSCRIPTION BASED
;    TITLE OF INVENTION:  AMPLIFICATION ASSAY FOR THE DETECTION AND QUANTITATION OF
;    TITLE OF INVENTION:  MACROPHAGE DERIVED CHEMOKINE RNA
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  AKZO NOBEL PATNT DEPARTMENT
;      STREET:  1300 PICCARD DRIVE, SUITE 206
;      CITY:  ROCKVILLE
;      STATE:  MARYLAND
;      COUNTRY:  US
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/005,165
;      FILING DATE:
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  KLESNER, SHARON N.
;      REGISTRATION NUMBER:  36,335
;      REFERENCE/DOCKET NUMBER:  MDC
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  301-948-7400
;      TELEFAX:  301-948-9751
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  24 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA
US-09-005-165-8

  Query Match             66.0%;  Score 13.2;  DB 3;  Length 24;
  Best Local Similarity   83.3%;  Pred. No. 5.9e+02;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 CCCAGGGCACTCATGGCT 18
           ||||||||||||  || |
Db      21 CCCAGGGCACTCTGGGAT 4


RESULT 3
US-08-289-222E-20
; Sequence 20, Application US/08289222E
; Patent No. 6120760
;  GENERAL INFORMATION:
;    APPLICANT:  HOTTEN, GERTRUD
;    APPLICANT:  NEIDHARDT, HELGE
;    APPLICANT:  BECHTOLD, ROLF
;    APPLICANT:  POHL, JENS
;    TITLE OF INVENTION:  GROWTH/DIFFERENTIATION FACTORS OF THE TGF-B
;    TITLE OF INVENTION:  FAMILY
;    NUMBER OF SEQUENCES:  53
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  NIKAIDO, MARMELSTEIN, MURRAY & ORAM
;      STREET:  655 FIFTEENTH STREET, N. W., G STREET LOBBY,
;      STREET:  SUITE 330
;      CITY:  WASHINGTON
;      STATE:  DC
;      COUNTRY:  USA
;      ZIP:  20005-5701
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/289,222E
;      FILING DATE:  25-AUG-1999
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/289,222
;      FILING DATE:  12-AUG-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  DE P 44 23 190.3
;      FILING DATE:  07-JUL-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  EPO 92102324.8
;      FILING DATE:  12-FEB-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/EP93/00350
;      FILING DATE:  12-FEB-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  KITTS, MONICA CHIN
;      REGISTRATION NUMBER:  36,105
;      REFERENCE/DOCKET NUMBER:  P564-9021
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202/638-5000
;      TELEFAX:  202/638-4810
;  INFORMATION FOR SEQ ID NO:  20:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  24 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  both
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA
US-08-289-222E-20

  Query Match             66.0%;  Score 13.2;  DB 3;  Length 24;
  Best Local Similarity   83.3%;  Pred. No. 5.9e+02;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 CCAGGGCACTCATGGCTA 19
           |||||||||| ||| | |
Db       2 CCAGGGCACTAATGTCAA 19


RESULT 4
US-09-054-526B-20
; Sequence 20, Application US/09054526B
; Patent No. 6197550
;  GENERAL INFORMATION:
;    APPLICANT:  H TTEN, GERTRUD
;    APPLICANT:  NEIDHARDT, HELGE
;    APPLICANT:  BECHTOLD, ROLF
;    APPLICANT:  POHL, JENS
;    TITLE OF INVENTION:  DNA SEQUENCES ENCODING NOVEL
;    TITLE OF INVENTION:  GROWTH/DIFFERENTATION FACTORS
;    NUMBER OF SEQUENCES:  53
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  NIKAIDO, MARMELSTEIN, MURRAY & ORAM LLP
;      STREET:  655 FIFTEENTH STREET, N. W., G STREET LOBBY,
;      STREET:  SUITE 330
```

```
;      CITY:  WASHINGTON
;      STATE:  DC
;      COUNTRY:  USA
;      ZIP:  20005-5701
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/054,526B
;      FILING DATE:  03-APR-1998
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/289,222
;      FILING DATE:  12-AUG-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  DE P 44 23 190.3
;      FILING DATE:  01-JUL-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  EPO 92102324.8
;      FILING DATE:  12-FEB-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/EP93/00350
;      FILING DATE:  12-FEB-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  KITTS, MONICA CHIN
;      REGISTRATION NUMBER:  36,105
;      REFERENCE/DOCKET NUMBER:  P564-8005
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202/638-5000
;      TELEFAX:  202/638-4810
;  INFORMATION FOR SEQ ID NO:  20:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  24 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  both
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA
US-09-054-526B-20

  Query Match             66.0%;  Score 13.2;  DB 4;  Length 24;
  Best Local Similarity   83.3%;  Pred. No. 5.9e+02;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 CCAGGGCACTCATGGCTA 19
           |||||||||| ||| | |
Db       2 CCAGGGCACTAATGTCAA 19


RESULT 5
US-08-242-663A-5
; Sequence 5, Application US/08242663A
; Patent No. 5492824
;  GENERAL INFORMATION:
;    APPLICANT:  Talanian, Robert V.
;    APPLICANT:  Dang, Leonard Luan Cao
;    APPLICANT:  Walker, Nigel Pelham Clinton
;    APPLICANT:  Ghayur, Tariq
;    TITLE OF INVENTION:  ICE AND ICE-LIKE COMPOSITIONS AND
;    TITLE OF INVENTION:  METHODS OF MAKING SAME
;    NUMBER OF SEQUENCES:  6
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Wolf, Greenfield & Sacks, P.C.
;      STREET:  600 Atlantic Avenue
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  USA
;      ZIP:  02210
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/242,663A
;      FILING DATE:  12-MAY-1994
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Janiuk, Anthony J.
;      REGISTRATION NUMBER:  29,809
;      REFERENCE/DOCKET NUMBER:  B0870/7001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-720-3500
;      TELEFAX:  617-720-2441
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  57 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
US-08-242-663A-5

  Query Match             64.0%;  Score 12.8;  DB 1;  Length 57;
  Best Local Similarity   87.5%;  Pred. No. 1e+03;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       5 GGGCACTCATGGCTAT 20
           ||| | ||||||||||
Db       2 GGGAATTCATGGCTAT 17


RESULT 6
US-08-573-890-3
; Sequence 3, Application US/08573890
; Patent No. 5869315
;  GENERAL INFORMATION:
;    APPLICANT:  Talanian, Robert V.
;    APPLICANT:  Mankovich, John A.
;    APPLICANT:  Ghayur, Tariq
;    APPLICANT:  Ferenz, Catherine R.
;    TITLE OF INVENTION:  Modified Human Interleukin-1b Converting
;    TITLE OF INVENTION:  Enzyme with Increased Stability
;    NUMBER OF SEQUENCES:  6
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  LAHIVE & COCKFIELD
;      STREET:  60 State Street, suite 510
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  USA
;      ZIP:  02109-1875
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  ASCII text
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/573,890
;      FILING DATE:
;      CLASSIFICATION:  536
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DECONTI, Giulio A., Jr.
;      REGISTRATION NUMBER:  31,503
;      REFERENCE/DOCKET NUMBER:  BBI-047
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617)227-7400
;      TELEFAX:  (617)227-5941
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  57 bases
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  oligonucleotide primer
```

```
US-08-573-890-3

  Query Match             64.0%;  Score 12.8;  DB 2;  Length 57;
  Best Local Similarity   87.5%;  Pred. No. 1e+03;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       5 GGGCACTCATGGCTAT 20
           ||| | ||||||||||
Db       2 GGGAATTCATGGCTAT 17


RESULT 7
US-09-248-179-3
; Sequence 3, Application US/09248179
; Patent No. 6242240
;  GENERAL INFORMATION:
;    APPLICANT:  Talanian, Robert V.
;    APPLICANT:  Mankovich, John A.
;    APPLICANT:  Ghayur, Tariq
;    APPLICANT:  Ferenz, Catherine R.
;    TITLE OF INVENTION:  Modified Human Interleukin-1b Converting
;    TITLE OF INVENTION:  Enzyme with Increased Stability
;    NUMBER OF SEQUENCES:  6
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  LAHIVE & COCKFIELD
;      STREET:  60 State Street, suite 510
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  USA
;      ZIP:  02109-1875
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  ASCII text
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/248,179
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/573,890
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DECONTI, Giulio A., Jr.
;      REGISTRATION NUMBER:  31,503
;      REFERENCE/DOCKET NUMBER:  BBI-047
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (617)227-7400
;      TELEFAX:  (617)227-5941
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  57 bases
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  oligonucleotide primer
US-09-248-179-3

  Query Match             64.0%;  Score 12.8;  DB 4;  Length 57;
  Best Local Similarity   87.5%;  Pred. No. 1e+03;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       5 GGGCACTCATGGCTAT 20
           ||| | ||||||||||
Db       2 GGGAATTCATGGCTAT 17


RESULT 8
PCT-US95-06132-5
; Sequence 5, Application PC/TUS9506132
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  ICE AND ICE-LIKE COMPOSITIONS AND
;    TITLE OF INVENTION:  METHODS OF MAKING SAME
;    NUMBER OF SEQUENCES:  6
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Wolf, Greenfield & Sacks, P.C.
;      STREET:  600 Atlantic Avenue
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  USA
;      ZIP:  02210
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/06132
;      FILING DATE:  FILED HEREWITH
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/242,663
;      FILING DATE:  12-MAY-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  JANIUK, Anthony J.
;      REGISTRATION NUMBER:  29,809
;      REFERENCE/DOCKET NUMBER:  B0870/7001WO
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-720-3500
;      TELEFAX:  617-720-2441
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  57 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
PCT-US95-06132-5

  Query Match             64.0%;  Score 12.8;  DB 5;  Length 57;
  Best Local Similarity   87.5%;  Pred. No. 1e+03;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       5 GGGCACTCATGGCTAT 20
           ||| | ||||||||||
Db       2 GGGAATTCATGGCTAT 17


RESULT 9
US-07-728-220C-14/c
; Sequence 14, Application US/07728220C
; Patent No. 6020168
;  GENERAL INFORMATION:
;    APPLICANT:  MATSUO, HISAYUKI
;    APPLICANT:  KANGAWA, KENJI
;    APPLICANT:  TANAKA, SHOJI
;    APPLICANT:  FUCHIMURA, KAYOKO
;    APPLICANT:  TAWARAGI, YASUNORI
;    TITLE OF INVENTION:  PORCINE CNP GENE AND PRECURSOR PROTEIN
;    NUMBER OF SEQUENCES:  22
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  CUSHMAN, DARBY & CUSHMAN
;      STREET:  ELEVENTH FLOOR, 1615 L STREET, N.W.
;      CITY:  WASHINGTON
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20036-5601
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/728,220C
```

```
;      FILING DATE:  19910712
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  SCOTT, WATSON T.
;      REGISTRATION NUMBER:  26,581
;      REFERENCE/DOCKET NUMBER:  WTS/9437/91816
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)861-3000
;      TELEFAX:  (202)822-0944
;      TELEX:  671 4627 CUSH
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  66 base pairs
;      TYPE:  NUCLEIC ACID
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-07-728-220C-14

  Query Match             64.0%;  Score 12.8;  DB 3;  Length 66;
  Best Local Similarity   87.5%;  Pred. No. 1e+03;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 CCCAGGGCACTCATGG 16
           |||||| | |||||||
Db      62 CCCAGGCCGCTCATGG 47


RESULT 10
US-08-352-179-10/c
; Sequence 10, Application US/08352179
; Patent No. 5670340
;  GENERAL INFORMATION:
;    APPLICANT:  YABUTA, Masayuki
;    APPLICANT:  SUZUKI, Yuji
;    APPLICANT:  OHSUYE, Kazuhiro
;    APPLICANT:  OSHIMA, Takehiro
;    APPLICANT:  ONAI, Seiko
;    APPLICANT:  MAGOTA, Koji
;    APPLICANT:  TANAKA, Shoji
;    TITLE OF INVENTION:  PROCESS FOR PRODUCING PEPTIDE
;    NUMBER OF SEQUENCES:  28
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Burns, Doane, Swecker and Mathis
;      STREET:  The George Mason Bldg., Washington & Prince
;      STREET:  Sts.
;      CITY:  Alexandria
;      STATE:  Virginia
;      COUNTRY:  United States
;      ZIP:  22313-1404
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/352,179
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/929,597
;      FILING DATE:  17-AUG-1992
;      APPLICATION NUMBER:  JP 3-320769
;      FILING DATE:  19-AUG-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 4-223520
;      FILING DATE:  31-JUL-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Crane-Feury, Sharon E
;      REGISTRATION NUMBER:  36,113
;      REFERENCE/DOCKET NUMBER:  001560-175
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 836-6620
;      TELEFAX:  (703) 836-2021
;  INFORMATION FOR SEQ ID NO:  10:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  69 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      CLONE:  plasmid pUCCNP1
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  1..69
US-08-352-179-10

  Query Match             64.0%;  Score 12.8;  DB 1;  Length 69;
  Best Local Similarity   87.5%;  Pred. No. 1e+03;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 CCCAGGGCACTCATGG 16
           |||||| | |||||||
Db      62 CCCAGGCCGCTCATGG 47


RESULT 11
US-08-823-516-105
; Sequence 105, Application US/08823516
; Patent No. 5994069
;  GENERAL INFORMATION:
;    APPLICANT:  Hall, Jeff G.
;    APPLICANT:  Lyamichev, Victor I.
;    APPLICANT:  Mast, Andrea L.
;    APPLICANT:  Brow, Mary Ann D.
;    TITLE OF INVENTION:  Detection Of Nucleic Acids By Multiple
;    TITLE OF INVENTION:  Sequential Invasive Cleavages
;    NUMBER OF SEQUENCES:  163
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Medlen & Carroll, LLP
;      STREET:  220 Montgomery Street, Suite 2200
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  United States Of America
;      ZIP:  94104
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/823,516
;      FILING DATE:  24-MAR-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US97/01072
;      FILING DATE:  21-JAN-1997
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/759,038
;      FILING DATE:  02-DEC-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/758,314
;      FILING DATE:  02-DEC-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/756,386
;      FILING DATE:  29-NOV-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/682,853
;      FILING DATE:  12-JUL-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/599,491
;      FILING DATE:  24-JAN-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Ingolia, Diane E.
;      REGISTRATION NUMBER:  40,027
;      REFERENCE/DOCKET NUMBER:  FORS-02736
```

```
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 705-8410
;      TELEFAX:  (415) 397-8338
;  INFORMATION FOR SEQ ID NO:  105:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  23 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  other nucleic acid
;      DESCRIPTION:  /desc = "DNA"
;    FEATURE:
;      NAME/KEY:  misc_difference
;      LOCATION:  replace(1, "")
;      OTHER INFORMATION:  /note= "The residue at this
;      OTHER INFORMATION:  position is a 2'deoxyguanosine 5'-O-(1-Thiomonophosphate)."
;    FEATURE:
;      NAME/KEY:  misc_difference
;      LOCATION:  replace(2, "")
;      OTHER INFORMATION:  /note= "The residue at this
;      OTHER INFORMATION:  position is a 2'deoxycytosine 5'-O-(1-Thiomonophosphate)."
;    FEATURE:
;      NAME/KEY:  misc_difference
;      LOCATION:  replace(3, "")
;      OTHER INFORMATION:  /note= "The residue at this
;      OTHER INFORMATION:  position is a 2'deoxythymidine 5'-O-(1-Thiomonophosphate)."
;    FEATURE:
;      NAME/KEY:  misc_difference
;      LOCATION:  replace(4, "")
;      OTHER INFORMATION:  /note= "The residue at this
;      OTHER INFORMATION:  position is a 2'deoxycytosine 5'-O-(1-Thiomonophosphate)."
;    FEATURE:
;      NAME/KEY:  misc_difference
;      LOCATION:  replace(5..6, "")
;      OTHER INFORMATION:  /note= "The residues at these
;      OTHER INFORMATION:  positions are 2'deoxyadenosine 5'-O-(1-Thiomonophosphate)."
;    FEATURE:
;      NAME/KEY:  misc_difference
;      LOCATION:  replace(7..8, "")
;      OTHER INFORMATION:  /note= "The residues at these
;      OTHER INFORMATION:  positions are 2'deoxyguanosine 5'-O-(1-Thiomonophosphate)."
;    FEATURE:
;      NAME/KEY:  misc_difference
;      LOCATION:  replace(9, "")
;      OTHER INFORMATION:  /note= "The residue at this
;      OTHER INFORMATION:  position is a 2'deoxycytosine 5'-O-(1-Thiomonophosphate)."
US-08-823-516-105

  Query Match             63.0%;  Score 12.6;  DB 2;  Length 23;
  Best Local Similarity   78.9%;  Pred. No. 1.1e+03;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 CCCAGGGCACTCATGGCTA 19
           | || ||||||| || |||
Db       2 CTCAAGGCACTCTTGCCTA 20


RESULT 12
US-08-759-038-90
; Sequence 90, Application US/08759038
; Patent No. 6090543
;  GENERAL INFORMATION:
;    APPLICANT:  Prudent, James R.
;    APPLICANT:  Hall, Jeff G.
;    APPLICANT:  Lyamichev, Victor I.
;    APPLICANT:  Brow, Mary Ann D.
;    APPLICANT:  Dahlberg, James E.
;    TITLE OF INVENTION:  Cleavage Of Nucleic Acids
;    NUMBER OF SEQUENCES:  134
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Medlen & Carroll, LLP
;      STREET:  220 Montgomery Street, Suite 2200
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  United States Of America
;      ZIP:  94104
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/759,038
;      FILING DATE:  02-DEC-1996
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/
;      FILING DATE:  29-NOV-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/682,853
;      FILING DATE:  12-JUL-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/599,491
;      FILING DATE:  24-JAN-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Ingolia, Diane E.
;      REGISTRATION NUMBER:  40,027
;      REFERENCE/DOCKET NUMBER:  FORS-02574
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 705-8410
;      TELEFAX:  (415) 397-8338
;  INFORMATION FOR SEQ ID NO:  90:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  23 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  other nucleic acid
;      DESCRIPTION:  /desc = "DNA"
;    FEATURE:
;      NAME/KEY:  misc_difference
;      LOCATION:  replace(1, "")
;      OTHER INFORMATION:  /note= "The residue at this
;      OTHER INFORMATION:  position is a 2'deoxyguanosine 5'-O-(1-Thiomonophosphate).
;    FEATURE:
;      NAME/KEY:  misc_difference
;      LOCATION:  replace(2, "")
;      OTHER INFORMATION:  /note= "The residue at this
;      OTHER INFORMATION:  position is a 2'deoxycytosine 5'-O-(1-Thiomonophosphate)."
;    FEATURE:
;      NAME/KEY:  misc_difference
;      LOCATION:  replace(3, "")
;      OTHER INFORMATION:  /note= "The residue at this
;      OTHER INFORMATION:  position is a 2'deoxythymidine 5'-O-(1-Thiomonophosphate).
;    FEATURE:
;      NAME/KEY:  misc_difference
;      LOCATION:  replace(4, "")
;      OTHER INFORMATION:  /note= "The residue at this
;      OTHER INFORMATION:  position is a 2'deoxycytosine 5'-O-(1-Thiomonophosphate)."
;    FEATURE:
;      NAME/KEY:  misc_difference
;      LOCATION:  replace(5..6, "")
;      OTHER INFORMATION:  /note= "The residues at these
;      OTHER INFORMATION:  positions are 2'deoxyadenosine 5'-O-(1-Thiomonophosphate).
;    FEATURE:
;      NAME/KEY:  misc_difference
;      LOCATION:  replace(7..8, "")
;      OTHER INFORMATION:  /note= "The residues at these
;      OTHER INFORMATION:  positions are 2'deoxyguanosine 5'-O-(1-Thiomonophosphate).
;    FEATURE:
;      NAME/KEY:  misc_difference
;      LOCATION:  replace(9, "")
;      OTHER INFORMATION:  /note= "The residue at this
;      OTHER INFORMATION:  position is a 2'deoxycytosine 5'-O-(1-Thiomonophosphate)."
US-08-759-038-90
```

```
  Query Match            63.0%;  Score 12.6;  DB 3;  Length 23;
  Best Local Similarity  78.9%;  Pred. No. 1.1e+03;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 CCCAGGGCACTCATGGCTA 19
          | || ||||||| || |||
Db      2 CTCAAGGCACTCTTGCCTA 20


RESULT 13
US-08-758-314-90
; Sequence 90, Application US/08758314
; Patent No. 6090606
;  GENERAL INFORMATION:
;    APPLICANT:  Kaiser, Michael W.
;    APPLICANT:  Lyamichev, Victor I.
;    APPLICANT:  Lyamichev, Natasha
;    TITLE OF INVENTION:  Improved Cleavage Agents
;    NUMBER OF SEQUENCES:  134
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Medlen & Carroll, LLP
;      STREET:  220 Montgomery Street, Suite 2200
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  United States Of America
;      ZIP:  94104
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/758,314
;      FILING DATE:  02-DEC-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/
;      FILING DATE:  29-NOV-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/682,853
;      FILING DATE:  12-JUL-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/599,491
;      FILING DATE:  24-JAN-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Ingolia, Diane E.
;      REGISTRATION NUMBER:  40,027
;      REFERENCE/DOCKET NUMBER:  FORS-02575
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 705-8410
;      TELEFAX:  (415) 397-8338
;  INFORMATION FOR SEQ ID NO:  90:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  23 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  other nucleic acid
;      DESCRIPTION:  /desc = "DNA"
;    FEATURE:
;      NAME/KEY:  misc_difference
;      LOCATION:  replace(1, "")
;      OTHER INFORMATION:  /note= "The residue at this
;      OTHER INFORMATION:  position is a 2'deoxyguanosine 5'-O-(1-Thiomonophosphate)."
;    FEATURE:
;      NAME/KEY:  misc_difference
;      LOCATION:  replace(2, "")
;      OTHER INFORMATION:  /note= "The residue at this
;      OTHER INFORMATION:  position is a 2'deoxycytosine 5'-O-(1-Thiomonophosphate)."
;    FEATURE:
;      NAME/KEY:  misc_difference
;      LOCATION:  replace(3, "")
;      OTHER INFORMATION:  /note= "The residue at this
;      OTHER INFORMATION:  position is a 2'deoxythymidine 5'-O-(1-Thiomonophosphate)."
;    FEATURE:
;      NAME/KEY:  misc_difference
;      LOCATION:  replace(4, "")
;      OTHER INFORMATION:  /note= "The residue at this
;      OTHER INFORMATION:  position is a 2'deoxycytosine 5'-O-(1-Thiomonophosphate)."
;    FEATURE:
;      NAME/KEY:  misc_difference
;      LOCATION:  replace(5..6, "")
;      OTHER INFORMATION:  /note= "The residues at these
;      OTHER INFORMATION:  positions are 2'deoxyadenosine 5'-O-(1-Thiomonophosphate).
;    FEATURE:
;      NAME/KEY:  misc_difference
;      LOCATION:  replace(7..8, "")
;      OTHER INFORMATION:  /note= "The residues at these
;      OTHER INFORMATION:  positions are 2'deoxyguanosine 5'-O-(1-Thiomonophosphate).
;    FEATURE:
;      NAME/KEY:  misc_difference
;      LOCATION:  replace(9, "")
;      OTHER INFORMATION:  /note= "The residue at this
;      OTHER INFORMATION:  position is a 2'deoxycytosine 5'-O-(1-Thiomonophosphate)."
US-08-758-314-90

  Query Match            63.0%;  Score 12.6;  DB 3;  Length 23;
  Best Local Similarity  78.9%;  Pred. No. 1.1e+03;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 CCCAGGGCACTCATGGCTA 19
          | || ||||||| || |||
Db      2 CTCAAGGCACTCTTGCCTA 20


RESULT 14
US-08-611-587-32
; Sequence 32, Application US/08611587
; Patent No. 6150091
;  GENERAL INFORMATION:
;    APPLICANT:  PANDOLFO, MASSIMO
;    APPLICANT:  MONTERMINI, LAURA
;    APPLICANT:  MOLTO, MARIA D.
;    APPLICANT:  Koenig, Michael
;    APPLICANT:  Campuzano, Victoria
;    APPLICANT:  Cossee, Mireille
;    TITLE OF INVENTION:  Direct Diagnosis of Friedreich Ataxia
;    NUMBER OF SEQUENCES:  33
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fulbright & Jaworski L.L.P. Patent Dept.
;      STREET:  1301 McKinney, Suite 5100
;      CITY:  Houston
;      STATE:  Texas
;      COUNTRY:  U.S.
;      ZIP:  77010
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/611,587
;      FILING DATE:  03-MAR-1996
;      CLASSIFICATION:  436
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Brashears-Macatee, Sarah J.
;      REGISTRATION NUMBER:  38,087
;      REFERENCE/DOCKET NUMBER:  D-5901
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  713-651-5620
;      TELEFAX:  713-651-5246
;      TELEX:  76-2829
;  INFORMATION FOR SEQ ID NO:  32:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  23 base pairs
;      TYPE:  nucleic acid
```

```
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  other nucleic acid
;       DESCRIPTION:  /desc = "oligo"
;     HYPOTHETICAL:  NO
;     ANTI-SENSE:  NO
;     POSITION IN GENOME:
;       UNITS:  bp
US-08-611-587-32

  Query Match             63.0%;  Score 12.6;  DB 3;  Length 23;
  Best Local Similarity   78.9%;  Pred. No. 1.1e+03;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 CCAGGGCACTCATGGCTAT 20
           ||||| || ||| |||| |
Db       5 CCAGGACAGTCAGGGCTTT 23


RESULT 15
US-08-526-136-5/c
; Sequence 5, Application US/08526136
; Patent No. 6107089
;  GENERAL INFORMATION:
;     APPLICANT:  Towle, Christine A. et al.
;     TITLE OF INVENTION:  ANNEXIN XI
;     NUMBER OF SEQUENCES:  36
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Fish & Richardson
;       STREET:  225 Franklin Street
;       CITY:  Boston
;       STATE:  Massachusetts
;       COUNTRY:  U.S.A.
;       ZIP:  02110-2804
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  3.5" Diskette, 1.44 Mb
;       COMPUTER:  IBM PS/2 Model 50Z or 55SX
;       OPERATING SYSTEM:  IBM P.C. DOS (Version 3.30)
;       SOFTWARE:  WordPerfect (Version 5.0)
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/526,136
;       FILING DATE:
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/214,036
;       FILING DATE:
;       APPLICATION NUMBER:  07/837,775
;       FILING DATE:  February 13, 1992
;       APPLICATION NUMBER:  07/764,465
;       FILING DATE:  September 23, 1991
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Clark, Paul T.
;       REGISTRATION NUMBER:  30,162
;       REFERENCE/DOCKET NUMBER:  00786/099001
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (617) 542-5070
;       TELEFAX:  (617) 542-8906
;       TELEX:  200154
;   INFORMATION FOR SEQ ID NO:  5:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  29
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
US-08-526-136-5

  Query Match             63.0%;  Score 12.6;  DB 3;  Length 29;
  Best Local Similarity   78.9%;  Pred. No. 1.2e+03;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 CCCAGGGCACTCATGGCTA 19
           ||  |||  ||||||||||
Db      24 CCTGGGGAGCTCATGGCTA 6
```

Search completed: November 23, 2002, 06:36:25
Job time : 22.55 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 05:54:41 ; Search time 17.25 Seconds
(without alignments)
439.108 Million cell updates/sec

Title: US-09-296-264-20
Perfect score: 20
Sequence: 1 cccagggcactcatggctat 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 335578 seqs, 189365133 residues

Total number of hits satisfying chosen parameters: 195614

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
c 1	14.4	72.0	39	10	US-09-879-919-16	Sequence 16, Appl
2	14	70.0	31	10	US-09-801-274-953	Sequence 953, App
3	13.6	68.0	82	10	US-09-878-574-8965	Sequence 8965, Ap
c 4	13.6	68.0	88	10	US-09-864-761-21663	Sequence 21663, A
5	13.4	67.0	78	10	US-09-938-700-25	Sequence 25, Appl
c 6	12.8	64.0	50	10	US-09-896-888A-10	Sequence 10, Appl
7	12.8	64.0	57	9	US-09-827-708A-3	Sequence 3, Appli
8	12.6	63.0	25	10	US-09-305-856B-93	Sequence 93, Appl
c 9	12.6	63.0	69	9	US-09-466-035-77	Sequence 77, Appl
c 10	12.6	63.0	69	10	US-09-912-679-77	Sequence 77, Appl
c 11	12.4	62.0	37	10	US-09-264-468B-16	Sequence 16, Appl
c 12	12.4	62.0	52	10	US-09-879-919-15	Sequence 15, Appl
13	12.4	62.0	90	10	US-09-864-761-22784	Sequence 22784, A
c 14	12.2	61.0	25	10	US-09-795-232-4	Sequence 4, Appli
c 15	12.2	61.0	26	10	US-09-854-864-29	Sequence 29, Appl
16	12.2	61.0	31	10	US-09-263-689-55	Sequence 55, Appl
c 17	12.2	61.0	32	9	US-10-067-618-12	Sequence 12, Appl
18	12.2	61.0	32	10	US-09-263-689-57	Sequence 57, Appl
19	12.2	61.0	33	9	US-09-319-264-3	Sequence 3, Appli
c 20	12.2	61.0	63	10	US-09-777-430A-59	Sequence 59, Appl
c 21	12.2	61.0	70	10	US-09-441-522-6	Sequence 6, Appli
c 22	12.2	61.0	70	10	US-09-441-522-7	Sequence 7, Appli
23	12.2	61.0	79	10	US-09-441-522-9	Sequence 9, Appli
24	12.2	61.0	91	10	US-09-764-877-3369	Sequence 3369, Ap
25	12	60.0	24	8	US-08-812-393A-3	Sequence 3, Appli
26	12	60.0	41	10	US-09-119-900-13	Sequence 13, Appl
c 27	12	60.0	79	10	US-09-775-174-2	Sequence 2, Appli
c 28	12	60.0	79	10	US-09-775-325-2	Sequence 2, Appli
c 29	12	60.0	87	10	US-09-864-761-19934	Sequence 19934, A
c 30	12	60.0	93	10	US-09-864-761-27086	Sequence 27086, A
31	12	60.0	99	10	US-09-960-352-7876	Sequence 7876, Ap
32	11.8	59.0	22	10	US-09-263-959-1086	Sequence 1086, Ap
33	11.8	59.0	27	10	US-09-916-230-34	Sequence 34, Appl
34	11.8	59.0	29	10	US-09-732-561-3	Sequence 3, Appli
35	11.8	59.0	33	10	US-09-961-933-2	Sequence 2, Appli
36	11.6	58.0	18	10	US-09-969-373-2803	Sequence 2803, Ap
37	11.6	58.0	28	10	US-09-275-805-3	Sequence 3, Appli
38	11.6	58.0	31	10	US-09-801-274-1254	Sequence 1254, Ap
c 39	11.6	58.0	31	10	US-09-430-221-15	Sequence 15, Appl
c 40	11.6	58.0	58	10	US-09-816-127-16	Sequence 16, Appl
41	11.6	58.0	66	10	US-09-923-246-32	Sequence 32, Appl
42	11.6	58.0	66	10	US-09-825-561A-24	Sequence 24, Appl
c 43	11.6	58.0	68	10	US-09-783-590-1368	Sequence 1368, Ap
c 44	11.6	58.0	69	9	US-09-832-659-25	Sequence 25, Appl
45	11.6	58.0	83	10	US-09-864-761-19080	Sequence 19080, A

ALIGNMENTS

RESULT 1
US-09-879-919-16/c
; Sequence 16, Application US/09879919
; Patent No. US20020064829A1
; GENERAL INFORMATION:
; APPLICANT: Yu, Guo-Liang, et al.
; TITLE OF INVENTION: Human Tumor Necrosis Factor Delta and Epsilon
; FILE REFERENCE: PF253P1
; CURRENT APPLICATION NUMBER: US/09/879,919
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: 60/293,499
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: 60/277,978
; PRIOR FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: 60/276,248
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/254,875
; PRIOR FILING DATE: 2000-12-13
; PRIOR APPLICATION NUMBER: 60/241,952
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/211,537
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 08/815,783
; PRIOR FILING DATE: 1997-03-12
; PRIOR APPLICATION NUMBER: 60/016,812
; PRIOR FILING DATE: 1996-03-14
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 16
; LENGTH: 39
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-879-919-16

Query Match 72.0%; Score 14.4; DB 10; Length 39;
Best Local Similarity 93.8%; Pred. No. 1.7e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy      1 CCCAGGGCACTCATGG 16
          |||||| |||||||||
Db     32 CCCAGGCCACTCATGG 17
```

```
RESULT 2
US-09-801-274-953
; Sequence 953, Application US/09801274
; Patent No. US20020032319A1
; GENERAL INFORMATION:
;  APPLICANT: Cargill, Michele
;  APPLICANT:  Ireland, James S.
;  APPLICANT:  Lander, Eric S.
;  TITLE OF INVENTION: HUMAN SINGLE NUCLEOTIDE POLYMORPHISMS
;  FILE REFERENCE: 2825.2009-001
;  CURRENT APPLICATION NUMBER: US/09/801,274
;  CURRENT FILING DATE:  2001-03-07
;  PRIOR APPLICATION NUMBER: US 60/187,510
;  PRIOR FILING DATE: 2000-03-07
;  PRIOR APPLICATION NUMBER: US 60/206,129
;  PRIOR FILING DATE: 2000-05-22
;  NUMBER OF SEQ ID NOS: 1802
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 953
;   LENGTH: 31
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-801-274-953

  Query Match             70.0%;  Score 14;  DB 10;  Length 31;
  Best Local Similarity   87.5%;  Pred. No. 2.7e+02;
  Matches   14;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       5 GGGCACTCATGGCTAT 20
            |||| |||||||:|||
Db       4 GGGCCCTCATGGMTAT 19


RESULT 3
US-09-878-574-8965
; Sequence 8965, Application US/09878574
; Patent No. US20020110548A1
; GENERAL INFORMATION:
;  APPLICANT: Byrum, Joseph R.
;  APPLICANT:  La Rosa, Thomas J.
;  APPLICANT:  Thompson, Michael D.
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
;  TITLE OF INVENTION:  Plants
;  FILE REFERENCE: 38-21(15401)B
;  CURRENT APPLICATION NUMBER: US/09/878,574
;  CURRENT FILING DATE:  2001-12-21
;  PRIOR APPLICATION NUMBER: 09/333,535
;  PRIOR FILING DATE: 1999-06-14
;  NUMBER OF SEQ ID NOS: 15775
; SEQ ID NO 8965
;   LENGTH: 82
;   TYPE: DNA
;   ORGANISM: Glycine max
;   OTHER INFORMATION: Clone ID: 701101903H1
US-09-878-574-8965

  Query Match             68.0%;  Score 13.6;  DB 10;  Length 82;
  Best Local Similarity   80.0%;  Pred. No. 4.6e+02;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 CCCAGGGCACTCATGGCTAT 20
            || |||  |||||||| |||
Db      20 CCAAGGATACTCATGGTTAT 39


RESULT 4
US-09-864-761-21663/c
; Sequence 21663, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 21663
;   LENGTH: 88
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AF152364.1
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.1
;   OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 0.99
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL =  4
;   OTHER INFORMATION: NT HIT: gi9506504, EVALUE 7.20e-01
;   OTHER INFORMATION: EST_HUMAN HIT: BE970003.1, EVALUE 4.20e+00
US-09-864-761-21663

  Query Match             68.0%;  Score 13.6;  DB 10;  Length 88;
  Best Local Similarity   80.0%;  Pred. No. 4.6e+02;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 CCCAGGGCACTCATGGCTAT 20
            ||||||||||||  ||  ||
Db      52 CCCAGGGCACTCCAGGTGAT 33


RESULT 5
US-09-938-700-25
; Sequence 25, Application US/09938700
; Patent No. US20020064525A1
; GENERAL INFORMATION:
;  APPLICANT: Morsey, et al.
```

```
;  TITLE OF INVENTION: Anti-IgE Vaccines
;  FILE REFERENCE: PC10761A
;  CURRENT APPLICATION NUMBER: US/09/938,700
;  CURRENT FILING DATE:  2001-08-24
;  NUMBER OF SEQ ID NOS: 28
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 25
;   LENGTH: 78
;   TYPE: DNA
;   ORGANISM: DOG CH3/CH4 NUCLEOTIDE SEQUENCE
US-09-938-700-25

  Query Match             67.0%;  Score 13.4;  DB 10;  Length 78;
  Best Local Similarity   93.3%;  Pred. No. 5.7e+02;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 CCCAGGGCACTCATG 15
          ||||||| ||||||
Db     25 CCCAGGGCCCTCATG 39


RESULT 6
US-09-896-888A-10/c
; Sequence 10, Application US/09896888A
; Patent No. US20020116723A1
; GENERAL INFORMATION:
;  APPLICANT: The University of British Columbia
;  TITLE OF INVENTION: Insect Expression Vectors
;  FILE REFERENCE: 80021-44
;  CURRENT APPLICATION NUMBER: US/09/896,888A
;  CURRENT FILING DATE:  2001-06-29
;  PRIOR APPLICATION NUMBER: US/09/048,911
;  PRIOR FILING DATE: 1998-03-26
;  PRIOR APPLICATION NUMBER: 60/049,946
;  PRIOR FILING DATE: 1997-03-27
;  NUMBER OF SEQ ID NOS: 50
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
;   LENGTH: 50
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Primer
US-09-896-888A-10

  Query Match             64.0%;  Score 12.8;  DB 10;  Length 50;
  Best Local Similarity   87.5%;  Pred. No. 1.1e+03;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      5 GGGCACTCATGGCTAT 20
          ||||| |||| |||||
Db     26 GGGCAGTCATCGCTAT 11


RESULT 7
US-09-827-708A-3
; Sequence 3, Application US/09827708A
; Patent No. US20020164762A1
;   GENERAL INFORMATION:
;        APPLICANT: Talanian, Robert V.
;                   Mankovich, John A.
;                   Ghayur, Tariq
;                   Ferenz, Catherine R.
;        TITLE OF INVENTION: Modified Human Interleukin-1b Converting
;                            Enzyme with Increased Stability
;        NUMBER OF SEQUENCES: 6
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: LAHIVE & COCKFIELD
;             STREET: 60 State Street, suite 510
;             CITY: Boston
;             STATE: Massachusetts
;             COUNTRY: USA
;             ZIP: 02109-1875
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: ASCII text
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/827,708A
;             FILING DATE: 06-Apr-2001
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 09/248,179
;             FILING DATE: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: DECONTI, Giulio A., Jr.
;             REGISTRATION NUMBER: 31,503
;             REFERENCE/DOCKET NUMBER: BBI-047
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (617)227-7400
;             TELEFAX: (617)227-5941
;   INFORMATION FOR SEQ ID NO: 3:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 57 bases
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: oligonucleotide primer
;        SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-827-708A-3

  Query Match             64.0%;  Score 12.8;  DB 9;  Length 57;
  Best Local Similarity   87.5%;  Pred. No. 1.1e+03;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      5 GGGCACTCATGGCTAT 20
          ||| | ||||||||||
Db      2 GGGAATTCATGGCTAT 17


RESULT 8
US-09-305-856B-93
; Sequence 93, Application US/09305856B
; Patent No. US20020061518A1
; GENERAL INFORMATION:
;  APPLICANT:  Penny, Laura
;  APPLICANT:  Galvin, Margaret
;  TITLE OF INVENTION:  Genotyping the Human
;  TITLE OF INVENTION:  UDP-Glucuronosyltransferase 1 (UGT1) Gene
;  FILE REFERENCE:  4389-7 (formerly SEQ-17CIP)
;  CURRENT APPLICATION NUMBER: US/09/305,856B
;  CURRENT FILING DATE:  1999-05-05
;  PRIOR APPLICATION NUMBER:  60/084,807
;  PRIOR FILING DATE:  1998-05-07
;  NUMBER OF SEQ ID NOS:  124
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 93
;   LENGTH: 25
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-305-856B-93

  Query Match             63.0%;  Score 12.6;  DB 10;  Length 25;
  Best Local Similarity   78.9%;  Pred. No. 1.3e+03;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 CCCAGGGCACTCATGGCTA 19
          |||||| || ||||| | |
Db      4 CCCAGGCCAATCATGCCCA 22


RESULT 9
US-09-466-035-77/c
; Sequence 77, Application US/09466035
; Patent No. US20020165172A1
```

```
;   GENERAL INFORMATION:
;        APPLICANT: SALLBERG, MATTI
;                   MILICH, DAVID R.
;                   LEE, WILLIAM T.L.
;        TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
;                            INTRACELLULAR DISEASES
;        NUMBER OF SEQUENCES: 86
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Robins & Pasternak LLP
;             STREET: 545 Middlefield Road, Suite 180
;             CITY: Menlo Park
;             STATE: California
;             COUNTRY: U.S.
;             ZIP: 94025
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.25
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/466,035
;             FILING DATE: 17-Dec-1999
;             CLASSIFICATION: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Pasternak, Dahna S.
;             REGISTRATION NUMBER: 41,411
;             REFERENCE/DOCKET NUMBER: 2300-1231.01
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 650-325-7812
;             TELEFAX: 650-325-7823
;             TELEX: <Unknown>
;   INFORMATION FOR SEQ ID NO: 77:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 69 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        SEQUENCE DESCRIPTION: SEQ ID NO: 77:
US-09-466-035-77

  Query Match             63.0%;  Score 12.6;  DB 9;  Length 69;
  Best Local Similarity   78.9%;  Pred. No. 1.4e+03;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      2 CCAGGGCACTCATGGCTAT 20
          | | | ||||||||||| |
Db     64 CAAAGTCACTCATGGCTTT 46


RESULT 10
US-09-912-679-77/c
; Sequence 77, Application US/09912679
; Patent No. US20020141974A1
;   GENERAL INFORMATION:
;        APPLICANT: Jolly, Douglas J.
;                   Chang, Stephen M.W.
;                   Lee, William T.L.
;                   Townsend, Kay
;                   O'Dea, Joanne
;        TITLE OF INVENTION: HEPATITIS THERAPEUTICS
;        NUMBER OF SEQUENCES: 84
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Seed and Berry
;             STREET: 6300 Columbia Center, 701 Fifth Avenue
;             CITY: Seattle
;             STATE: Washington
;             COUNTRY: U.S.
;             ZIP: 98104
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.25
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/912,679
;             FILING DATE: 07-Jun-1995
;             CLASSIFICATION: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: McMasters, David D.
;             REGISTRATION NUMBER: 33,963
;             REFERENCE/DOCKET NUMBER: 930049.407C5
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 206-622-4900
;             TELEFAX: 206-682-6031
;             TELEX: 3723836
;   INFORMATION FOR SEQ ID NO: 77:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 69 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        SEQUENCE DESCRIPTION: SEQ ID NO: 77:
US-09-912-679-77

  Query Match             63.0%;  Score 12.6;  DB 10;  Length 69;
  Best Local Similarity   78.9%;  Pred. No. 1.4e+03;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      2 CCAGGGCACTCATGGCTAT 20
          | | | ||||||||||| |
Db     64 CAAAGTCACTCATGGCTTT 46


RESULT 11
US-09-264-468B-16/c
; Sequence 16, Application US/09264468B
; Patent No. US20020106775A1
; GENERAL INFORMATION:
;  APPLICANT: Wang, Jieyi
;  APPLICANT:  Nienaber, Vicki L.
;  APPLICANT:  Henkin, Jack
;  APPLICANT:  Smith, Richard A.
;  APPLICANT:  Walter, Karl A.
;  APPLICANT:  Severin, Jean M.
;  APPLICANT:  Edalji, Rohinton
;  APPLICANT:  Johnson Jr., Robert W.
;  APPLICANT:  Holzman, Thomas F.
;  TITLE OF INVENTION: HIGHLY CRYSTALLINE UROKINASE
;  FILE REFERENCE: 6310.US.P1
;  CURRENT APPLICATION NUMBER: US/09/264,468B
;  CURRENT FILING DATE:  1999-03-05
;  PRIOR APPLICATION NUMBER: US 09/036,361
;  PRIOR FILING DATE: 1998-03-06
;  NUMBER OF SEQ ID NOS: 23
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 16
;   LENGTH: 37
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: PCR primer
US-09-264-468B-16

  Query Match             62.0%;  Score 12.4;  DB 10;  Length 37;
  Best Local Similarity   92.9%;  Pred. No. 1.7e+03;
  Matches   13;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      3 CAGGGCACTCATGG 16
          |||||| |||||||
Db     35 CAGGGCTCTCATGG 22


RESULT 12
US-09-879-919-15/c
; Sequence 15, Application US/09879919
; Patent No. US20020064829A1
```

```
; GENERAL INFORMATION:
;  APPLICANT: Yu, Guo-Liang, et al.
;  TITLE OF INVENTION: Human Tumor Necrosis Factor Delta and Epsilon
;  FILE REFERENCE: PF253P1
;  CURRENT APPLICATION NUMBER: US/09/879,919
;  CURRENT FILING DATE:  2001-06-14
;  PRIOR APPLICATION NUMBER: 60/293,499
;  PRIOR FILING DATE: 2001-05-25
;  PRIOR APPLICATION NUMBER: 60/277,978
;  PRIOR FILING DATE: 2001-03-23
;  PRIOR APPLICATION NUMBER: 60/276,248
;  PRIOR FILING DATE: 2001-03-16
;  PRIOR APPLICATION NUMBER: 60/254,875
;  PRIOR FILING DATE: 2000-12-13
;  PRIOR APPLICATION NUMBER: 60/241,952
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/211,537
;  PRIOR FILING DATE: 2000-06-15
;  PRIOR APPLICATION NUMBER: 08/815,783
;  PRIOR FILING DATE: 1997-03-12
;  PRIOR APPLICATION NUMBER: 60/016,812
;  PRIOR FILING DATE: 1996-03-14
;  NUMBER OF SEQ ID NOS: 26
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
;   LENGTH: 52
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-879-919-15

  Query Match             62.0%;  Score 12.4;  DB 10;  Length 52;
  Best Local Similarity   92.9%;  Pred. No. 1.7e+03;
  Matches   13;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 CCCAGGGCACTCAT 14
           |||||| |||||||
Db      14 CCCAGGCCACTCAT 1


RESULT 13
US-09-864-761-22784
; Sequence 22784, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 22784
;   LENGTH: 90
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AC004615.1
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 5.1
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 5.6
;   OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 5.3
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 5.6
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 4.7
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 5.4
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 5.1
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 4.9
;   OTHER INFORMATION: EST_HUMAN HIT: AA815175.1, EVALUE 3.00e-07
;   OTHER INFORMATION: NT HIT: AF111168.2, EVALUE 5.00e-05
US-09-864-761-22784

  Query Match             62.0%;  Score 12.4;  DB 10;  Length 90;
  Best Local Similarity   92.9%;  Pred. No. 1.8e+03;
  Matches   13;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       7 GCACTCATGGCTAT 20
           | ||||||||||||
Db       1 GAACTCATGGCTAT 14


RESULT 14
US-09-795-232-4/c
; Sequence 4, Application US/09795232
; Patent No. US20010012627A1
; GENERAL INFORMATION:
;  APPLICANT:  Anthony M. Brown
;  APPLICANT:  Conrad Gerald Chapman
;  APPLICANT:  Israel Simon Gloger
;  APPLICANT:  Joanne Rachel Evans
;  APPLICANT:  William Cairns
;  APPLICANT:  Hugh Jonathan Herdon
;  TITLE OF INVENTION:  NOVEL COMPOUNDS
;  FILE REFERENCE:  GP-30176-D1
;  CURRENT APPLICATION NUMBER: US/09/795,232
;  CURRENT FILING DATE:  2001-02-28
;  PRIOR APPLICATION NUMBER:  09/182,728
;  PRIOR FILING DATE:  1998-10-29
;  PRIOR APPLICATION NUMBER:  9818890.7
;  PRIOR FILING DATE:  1998-08-28
;  NUMBER OF SEQ ID NOS:  6
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 4
;   LENGTH: 25
;   TYPE: DNA
;   ORGANISM: HOMO SAPIENS
US-09-795-232-4

  Query Match             61.0%;  Score 12.2;  DB 10;  Length 25;
```

```
  Best Local Similarity   82.4%;  Pred. No. 2e+03;
  Matches   14;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       3 CAGGGCACTCATGGCTA 19
           | |||||||||  ||||
Db      21 CTGGGCACTCAGTGCTA 5


RESULT 15
US-09-854-864-29/c
; Sequence 29, Application US/09854864
; Patent No. US20020081296A1
; GENERAL INFORMATION:
;  APPLICANT: THEILL, LARS EYDE
;  APPLICANT:  YU, GANG
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS OF MATTER CONCERNING APRIL/G70, BCMA,
;  TITLE OF INVENTION:  BLYS/AGP-3, AND TACI
;  FILE REFERENCE: A-686B
;  CURRENT APPLICATION NUMBER: US/09/854,864
;  CURRENT FILING DATE:  2001-09-11
;  PRIOR APPLICATION NUMBER: US 60/204,039
;  PRIOR FILING DATE: 2000-05-12
;  PRIOR APPLICATION NUMBER: US 60/214,591
;  PRIOR FILING DATE: 2000-06-27
;  NUMBER OF SEQ ID NOS: 31
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 29
;   LENGTH: 26
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-854-864-29

  Query Match            61.0%;  Score 12.2;  DB 10;  Length 26;
  Best Local Similarity   82.4%;  Pred. No. 2e+03;
  Matches   14;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 CCAGGGCACTCATGGCT 18
           ||||| |||||||  ||
Db      26 CCAGGCCACTCATTACT 10


Search completed: November 23, 2002, 06:42:15
Job time : 18.25 secs
```

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 25, 2002, 09:10:06 ; Search time 755.55 Seconds
(without alignments)
428.707 Million cell updates/sec

Title: US-09-296-264-20
Perfect score: 20
Sequence: 1 cccagggcactcatggctat 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 357874

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	14.2	71.0	54	9	AA623642	AA623642 vq68d05.s
c	2	13.8	69.0	78	9	AA958370	AA958370 ua10f08.r
	3	13.6	68.0	70	17	AF107425	AF107425 AF107425
c	4	13.6	68.0	78	17	AL761883	AL761883 Arabidops
c	5	13.6	68.0	97	9	AA023719	AA023719 mh78b11.r
c	6	13.4	67.0	52	17	AZ772251	AZ772251 1M0583D04
c	7	13.4	67.0	96	17	AZ786158	AZ786158 2M0031E01
	8	13.2	66.0	34	10	AV963376	AV963376 AV963376
c	9	13.2	66.0	38	17	AZ596225	AZ596225 1M0409D19
c	10	13.2	66.0	50	10	BE042337	BE042337 hk35g08.y
	11	13.2	66.0	54	9	AU259826	AU259826 AU259826
	12	13.2	66.0	56	17	AZ770399	AZ770399 1M0571O19
c	13	13.2	66.0	60	9	AA426872	AA426872 vf21c02.s
	14	13.2	66.0	62	17	AZ592652	AZ592652 1M0403A13
	15	13.2	66.0	75	9	AU260179	AU260179 AU260179
c	16	13.2	66.0	81	17	BH218228	BH218228 1006078B0
	17	13.2	66.0	85	9	AA575523	AA575523 vm03b04.r
	18	13.2	66.0	90	12	BG314789	BG314789 OP1.0.254
	19	13.2	66.0	91	17	AZ767595	AZ767595 1M0566D24
	20	13.2	66.0	95	17	AZ785550	AZ785550 2M0029F02
	21	13.2	66.0	100	9	AI094970	AI094970 qa09h07.x
c	22	12.8	64.0	24	17	AZ428700	AZ428700 1M0212L16
	23	12.8	64.0	35	17	AZ621958	AZ621958 1M0455N23
c	24	12.8	64.0	50	9	AU105199	AU105199 AU105199
c	25	12.8	64.0	50	9	AU105201	AU105201 AU105201
c	26	12.8	64.0	50	9	AU105204	AU105204 AU105204
c	27	12.8	64.0	50	9	AU105205	AU105205 AU105205
c	28	12.8	64.0	50	9	AU105206	AU105206 AU105206
c	29	12.8	64.0	50	9	AU105207	AU105207 AU105207
c	30	12.8	64.0	96	9	AA895160	AA895160 vx47c11.r
c	31	12.8	64.0	98	10	AV919935	AV919935 AV919935
	32	12.8	64.0	100	17	BH771280	BH771280 hg30b12.b
c	33	12.8	64.0	100	17	BH866565	BH866565 hg30b12.y
c	34	12.6	63.0	38	17	AZ642713	AZ642713 1M0505D21
	35	12.6	63.0	40	17	AZ622990	AZ622990 1M0460N12
	36	12.6	63.0	41	14	H28625	H28625 yl52a06.r1
c	37	12.6	63.0	45	9	AI538057	AI538057 to83b07.x
c	38	12.6	63.0	47	17	AZ779504	AZ779504 2M0016E04
	39	12.6	63.0	52	9	AI310821	AI310821 qo93g12.x
c	40	12.6	63.0	52	14	W90057	W90057 zh69h10.r1
c	41	12.6	63.0	53	17	BH411420	BH411420 1007022F0
c	42	12.6	63.0	54	17	AZ346887	AZ346887 1M0082P17
	43	12.6	63.0	56	9	AA414801	AA414801 vc72g12.s
	44	12.6	63.0	58	9	AA619745	AA619745 vl60e01.s
	45	12.6	63.0	61	9	AA414777	AA414777 vc69e10.s

ALIGNMENTS

RESULT 1
AA623642
LOCUS AA623642 54 bp mRNA linear EST 14-OCT-1997
DEFINITION vq68d05.s1 Knowles Solter mouse 2 cell Mus musculus cDNA clone IMAGE:1107465 5', mRNA sequence.
ACCESSION AA623642
VERSION AA623642.1 GI:2527518
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 54)
AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.
TITLE The WashU-HHMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Marra M/Mouse EST Project
WashU-HHMI Mouse EST Project
Washington University School of MedicineP
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.

MGI:605633.
FEATURES Location/Qualifiers
source 1. .54
/organism="Mus musculus"
/strain="C57BL/6J x DBA/2J F1"
/db_xref="taxon:10090"
/clone="IMAGE:1107465"
/clone_lib="Knowles Solter mouse 2 cell"
/tissue_type="embryo"
/dev_stage="2-cell"
/lab_host="DH10B"
/note="Organ: embryo; Vector: pBluescribe (modified); Site_1: MluI; Site_2: SalI; Cloned unidirectionally from mRNA prepared from 13,500 2-cell stage embryos. Primer: SalI(dT): 5'-CGGTCGACCGTCGACCGTTTTTTTTTTTTTT-3'. cDNAs were cloned into the MluI/SalI sites of a modified pBluescribe vector using commercial linkers (NEB). Average insert size: 1.2 kb."
BASE COUNT 19 a 14 c 12 g 9 t
ORIGIN

Query Match 71.0%; Score 14.2; DB 9; Length 54;
Best Local Similarity 84.2%; Pred. No. 7e+03;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 CCAGGGCACTCATGGCTAT 20
||||| || |||||||||
Db 9 CCAGGACAGCCATGGCTAT 27

RESULT 2
AA958370/c
LOCUS AA958370 78 bp mRNA linear EST 08-MAY-1998
DEFINITION ua10f08.r1 Soares_mammary_gland_NbMMG Mus musculus cDNA clone IMAGE:1346343 5', mRNA sequence.
ACCESSION AA958370
VERSION AA958370.1 GI:3124600
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 78)
AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.
TITLE The WashU-HHMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Marra M/Mouse EST Project
WashU-HHMI Mouse EST Project
Washington University School of MedicineP
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:695135
Seq primer: -28m13 rev2 ET from Amersham
High quality sequence stop: 67.
FEATURES Location/Qualifiers
source 1. .78
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:1346343"
/clone_lib="Soares_mammary_gland_NbMMG"
/sex="male"
/tissue_type="mammary gland"
/dev_stage="4 weeks"
/lab_host="DH10B"
/note="Organ: mammary gland; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGGAGCGGCCGCGAATGGTTTTTTTTTTTTTTTTTTTTT T 3']; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. RNA provided by Dr. Minoru Ko, Wayne State Univ. Library constructed and normalized by Bento Soares and M.Fatima Bonaldo."
BASE COUNT 23 a 18 c 15 g 22 t
ORIGIN

Query Match 69.0%; Score 13.8; DB 9; Length 78;
Best Local Similarity 88.2%; Pred. No. 1.3e+04;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CCCAGGGCACTCATGGC 17
||||||| |||| ||||
Db 71 CCCAGGGAACTCTTGGC 55

RESULT 3
AF107425
LOCUS AF107425 70 bp DNA linear GSS 21-FEB-2001
DEFINITION AF107425 Human Homo sapiens genomic, DNA sequence.
ACCESSION AF107425
VERSION AF107425.1 GI:5281156
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 70)
AUTHORS Heus,H.C., Hing,A., van Baren,M.J., Joosse,M., Breedveld,G., Wang,J.C., Burgess,A., Donnis-Keller,H., Berglund,C., Zguricas,J., Scherer,S.W., Rommens,J.M., Oostra,B.A. and Heutink,P.
TITLE A physical and transcriptional map of the preaxial polydactyly locus on chromosome 7q36
JOURNAL Genomics 57 (3), 342-351 (1999)
MEDLINE 99263496
COMMENT Contact: Heus HC
Clinical Genetics
Erasmus University Rotterdam
Postbus 1738, Rotterdam, 3000 DR, The Netherlands
putative exon 22 of the preaxial polydactyly locus
Class: unknown.
FEATURES Location/Qualifiers
source 1. .70
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="7q36"
/clone_lib="Human"
/note="Vector: pMOS Blue"
BASE COUNT 14 a 18 c 22 g 16 t
ORIGIN

Query Match 68.0%; Score 13.6; DB 17; Length 70;
Best Local Similarity 80.0%; Pred. No. 1.5e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 CCCAGGGCACTCATGGCTAT 20
|||||| ||| |||||| |
Db 45 CCCAGGACACCAATGGCTGT 64

RESULT 4
AL761883/c
LOCUS AL761883 78 bp DNA linear GSS 18-JUN-2002
DEFINITION Arabidopsis thaliana T-DNA flanking sequence GK-229B08-014264, genomic survey sequence.
ACCESSION AL761883

```
VERSION     AL761883.1  GI:21505653
KEYWORDS    GSS.
SOURCE      thale cress.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1
  AUTHORS   Strizhov,N., Li,Y., Rosso,M., Viehoever,P., Dekker,K., Saedler,H.
            and Weisshaar,B.
  TITLE     A pipeline for automated high-throughput generation of FSTs
            (flanking sequence tags) from Arabidopsis thaliana T-DNA
            transformed lines
  JOURNAL   Unpublished
REFERENCE   2
  AUTHORS   Rosso,M., Strizhov,N., Li,Y., Reiss,B., Dekker,K. and Weisshaar,B.
  TITLE     A new Arabidopsis thaliana T-DNA mutagenised population (GABI-Kat)
            for flanking sequence tag based reverse genetics
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 78)
  AUTHORS   Strizhov,N., Li,Y., Rosso,M. and Weisshaar,B.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JUN-2002) Weisshaar B., Max-Planck-Institut fuer
            Zuechtungsforschung, Carl-von-Linne-Weg 10, Koeln, 50829, Germany
COMMENT     This sequence is recovered from the left border of the T-DNA. It
            indicates an insertion within the locus defined by clone f23f1. The
            sequences are generated at the MPI for Plant Breeding Research in
            the context of the GABI-Kat project. GABI-Kat is part of the German
            Plant Genomics program designated 'GABI'. Information on line
            availability can be found at:
            http://www.mpiz-koeln.mpg.de/GABI-Kat/.
FEATURES             Location/Qualifiers
     source          1. .78
                     /organism="Arabidopsis thaliana"
                     /strain="Columbia 0"
                     /db_xref="taxon:3702"
                     /clone="GK-229B08-014264"
                     /clone_lib="Arabidopsis thaliana T-DNA insertion lines"
                     /note="PCR was performed on DNA from Arabidopsis thaliana
                     plants (T1) which were transformed with the T-DNA from
                     vector pAC161. The lines contain one or more T-DNA
                     insertions. The DNA fragment(s) resulting from the PCR
                     were directly sequenced to determine the genomic sequence
                     flanking the insertion. Sequences displaying significant
                     similarity to the A. thalaina nuclear genome sequence were
                     processed for submission. T-DNA derived sequences were
                     removed"
BASE COUNT       23 a     11 c     15 g     23 t      6 others
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 17;  Length 78;
  Best Local Similarity   80.0%;  Pred. No. 1.6e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 CCCAGGGCACTCATGGCTAT 20
           ||||  ||||||||| | ||
Db      68 CCCATAGCACTCATGACAAT 49


RESULT 5
AA023719/c
LOCUS       AA023719                 97 bp    mRNA    linear   EST 21-JAN-1997
DEFINITION  mh78b11.r1 Soares mouse placenta 4NbMP13.5 14.5 Mus musculus cDNA
            clone IMAGE:457053 5' similar to SW:RAC1_HUMAN P15154 RAS-RELATED
            C3 BOTULINUM TOXIN SUBSTRATE 1 ;, mRNA sequence.
ACCESSION   AA023719
VERSION     AA023719.1  GI:1487634
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 97)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:273941
            Trace considered overall poor quality
            Possible reversed clone: similarity on wrong strand
            Seq primer: -28M13 rev2 from Amersham
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .97
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:457053"
                     /clone_lib="Soares mouse placenta 4NbMP13.5 14.5"
                     /sex="unknown"
                     /tissue_type="placenta"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /note="Organ: placenta; Vector: pT7T3D-Pac (Pharmacia)
                     with a modified polylinker; Site_1: Not I; Site_2: Eco RI;
                     1st strand cDNA was primed with a Not I - oligo(dT) primer
                     [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCGGAAATTTTTTTTTTTTTTTTTTTTTT
                     T 3']; double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. Library
                     went through one round of normalization, and was
                     constructed by Bento Soares and M.Fatima Bonaldo."
BASE COUNT       26 a     23 c     29 g     19 t
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 9;  Length 97;
  Best Local Similarity   80.0%;  Pred. No. 1.7e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 CCCAGGGCACTCATGGCTAT 20
           || |  ||||||| ||||||
Db      93 CCTATAGCACTCACGGCTAT 74


RESULT 6
AZ772251/c
LOCUS       AZ772251                 52 bp    DNA     linear   GSS 16-FEB-2001
DEFINITION  1M0583D04F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0583D04 F, DNA sequence.
ACCESSION   AZ772251
VERSION     AZ772251.1  GI:12895361
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 52)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
```

```
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0583  row: D  column: 04
            Seq primer: CGTTGTAAAACGACGGCCAGT
            Class: plasmid ends
            High quality sequence stop: 52.
FEATURES             Location/Qualifiers
     source          1. .52
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC1M0583D04"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
BASE COUNT       10 a     15 c      8 g     19 t
ORIGIN

  Query Match             67.0%;  Score 13.4;  DB 17;  Length 52;
  Best Local Similarity   93.3%;  Pred. No. 1.6e+04;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      6 GGCACTCATGGCTAT 20
          || |||||||||||||
Db     35 GGGACTCATGGCTAT 21


RESULT 7
AZ786158/c
LOCUS       AZ786158                 96 bp    DNA     linear   GSS 16-FEB-2001
DEFINITION  2M0031E01R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC2M0031E01 R, DNA sequence.
ACCESSION   AZ786158
VERSION     AZ786158.1  GI:12923638
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 96)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0031  row: E  column: 01
            Seq primer: CACACAGGAAACAGCTATGACC
            Class: plasmid ends
            High quality sequence stop: 96.
FEATURES             Location/Qualifiers
     source          1. .96
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC2M0031E01"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
BASE COUNT       39 a     25 c     19 g     13 t
ORIGIN

  Query Match             67.0%;  Score 13.4;  DB 17;  Length 96;
  Best Local Similarity   93.3%;  Pred. No. 2.1e+04;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      6 GGCACTCATGGCTAT 20
          |||||||| ||||||
Db     40 GGCACTCAGGGCTAT 26


RESULT 8
AV963376
LOCUS       AV963376                 34 bp    mRNA    linear   EST 14-MAR-2002
DEFINITION  AV963376 Nori Satoh unpublished cDNA library, egg Ciona
            intestinalis cDNA clone cieg21i24 5', mRNA sequence.
ACCESSION   AV963376
VERSION     AV963376.1  GI:19451675
KEYWORDS    EST.
SOURCE      Ciona intestinalis.
  ORGANISM  Ciona intestinalis
            Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
            Phlebobranchia; Cionidae; Ciona.
REFERENCE   1  (bases 1 to 34)
  AUTHORS   Satoh,N., Satou,Y., Kohara,Y. and Shin-i,T.
  TITLE     Expressed genes in Ciona intestinalis
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Nori Satoh
            Department of Zoology
            Kyoto University
```

```
           Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
           Tel: 81-75-753-4081
           Fax: 81-75-705-1113
           Email: satoh@ascidian.zool.kyoto-u.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .34
                     /organism="Ciona intestinalis"
                     /db_xref="taxon:7719"
                     /clone="cieg21i24"
                     /clone_lib="Nori Satoh unpublished cDNA library, egg"
                     /tissue_type="whole animal"
                     /dev_stage="egg"
                     /note="Vector: pBluescript SK"
BASE COUNT        9 a     12 c      5 g      8 t
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 10;  Length 34;
  Best Local Similarity   83.3%;  Pred. No. 1.7e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 CCCAGGGCACTCATGGCT 18
          |||| | || ||||||||
Db      5 CCCACGCCAATCATGGCT 22


RESULT 9
AZ596225/c
LOCUS       AZ596225                38 bp    DNA     linear   GSS 13-DEC-2000
DEFINITION  1M0409D19F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0409D19 F, DNA sequence.
ACCESSION   AZ596225
VERSION     AZ596225.1  GI:11718415
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 38)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0409  row: D  column: 19
            Seq primer: CGTTGTAAAACGACGGCCAGT
            Class: plasmid ends
            High quality sequence stop: 38.
FEATURES             Location/Qualifiers
     source          1. .38
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC1M0409D19"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
BASE COUNT        9 a      4 c     16 g      9 t
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 17;  Length 38;
  Best Local Similarity   83.3%;  Pred. No. 1.8e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 CCAGGGCACTCATGGCTA 19
          || | ||| |||||||||
Db     24 CCTGTGCAATCATGGCTA 7


RESULT 10
BE042337/c
LOCUS       BE042337                50 bp    mRNA    linear   EST 08-JUN-2000
DEFINITION  hk35g08.y1 NCI_CGAP_Ov34 Homo sapiens cDNA clone IMAGE:2998718 5',
            mRNA sequence.
ACCESSION   BE042337
VERSION     BE042337.1  GI:8359390
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            cDNA Library Preparation: David B. Krizman, Ph.D.
             cDNA Library Arrayed by: I.M.A.G.E. Consortium, LLNL
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL, send email to:
            info@image.llnl.gov
            Seq primer: -40RP from Gibco.
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2998718"
                     /clone_lib="NCI_CGAP_Ov34"
                     /sex="female"
                     /tissue_type="borderline ovarian carcinoma"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /note="Organ: ovary; Vector: pAMP1; mRNA made from
                     borderline ovarian carcinoma, cDNA made by oligo-dT
                     priming.  Directionally cloned.  Size-selected on agarose
                     gel, average insert size 500 bp.  Primary library,
                     non-amplified.  "
BASE COUNT       17 a      6 c     16 g     11 t
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 10;  Length 50;
  Best Local Similarity   83.3%;  Pred. No. 2e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;
```

```
Qy      2 CCAGGGCACTCATGGCTA 19
          |||| |||| ||| ||||
Db     49 CCAGCGCACACATTGCTA 32


RESULT 11
AU259826
LOCUS       AU259826                 54 bp    mRNA    linear   EST 25-APR-2002
DEFINITION  AU259826 3'-directed mouse cDNA library Mus musculus cDNA clone
            BED0015801 3', mRNA sequence.
ACCESSION   AU259826
VERSION     AU259826.1  GI:20326719
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 54)
  AUTHORS   Kato,K. and Matoba,R.
  TITLE     Generation of expressed sequence tags from mouse brain
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Kikuya Kato
            Graduate School of Biological Sciences
            Nara Institute of Science and Technology
            8916-5 Takayama, Ikoma, Nara 630-0101, Japan
            Tel: 81-743-72-5581
            Fax: 81-743-72-5589
            Email: kkato@bs.aist-nara.ac.jp,
            URL:http://love2.aist-nara.ac.jp/BED/index.html.
FEATURES             Location/Qualifiers
     source          1. .54
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /clone="BED0015801"
                     /clone_lib="3'-directed mouse cDNA library"
                     /tissue_type="brain"
                     /note="Vector: pGEM-T-easy"
BASE COUNT       14 a     13 c     11 g     16 t
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 9;  Length 54;
  Best Local Similarity   83.3%;  Pred. No. 2.1e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      3 CAGGGCACTCATGGCTAT 20
          |||| || || |||||||
Db     16 CAGGACAGTCCTGGCTAT 33


RESULT 12
AZ770399
LOCUS       AZ770399                 56 bp    DNA     linear   GSS 16-FEB-2001
DEFINITION  1M0571O19R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0571O19 R, DNA sequence.
ACCESSION   AZ770399
VERSION     AZ770399.1  GI:12891544
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 56)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0571  row: O  column: 19
            Seq primer: CACACAGGAAACAGCTATGACC
            Class: plasmid ends
            High quality sequence stop: 56.
FEATURES             Location/Qualifiers
     source          1. .56
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC1M0571O19"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
BASE COUNT       28 a      8 c     12 g      8 t
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 17;  Length 56;
  Best Local Similarity   83.3%;  Pred. No. 2.1e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      3 CAGGGCACTCATGGCTAT 20
          |||||||| ||||  |||
Db     34 CAGGGCACACATGAATAT 51


RESULT 13
AA426872/c
LOCUS       AA426872                 60 bp    mRNA    linear   EST 16-OCT-1997
DEFINITION  vf21c02.s1 Knowles Solter mouse unfertilized egg Mus musculus cDNA
            clone IMAGE:836354 5' similar to SW:RAPA_DISOM P22123 RAS-RELATED
            PROTEIN O-KREV. ;, mRNA sequence.
ACCESSION   AA426872
VERSION     AA426872.1  GI:2109650
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 60)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
```

```
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:496570
            Possible reversed clone: similarity on wrong strand
            Seq primer: -40m13 fwd. ET from Amersham.
FEATURES              Location/Qualifiers
     source           1. .60
                      /organism="Mus musculus"
                      /strain="C57BL/6J x DBA/2J F1"
                      /db_xref="taxon:10090"
                      /clone="IMAGE:836354"
                      /clone_lib="Knowles Solter mouse unfertilized egg"
                      /tissue_type="unfertilized egg"
                      /lab_host="DH10B"
                      /note="Organ: unfertilized egg; Vector: pBluescribe
                      (modified); Site_1: MluI; Site_2: SalI; Cloned
                      unidirectionally from mRNA prepared from 5000 unfertilized
                      eggs.  Primer: SalI(dT):
                      5'-CGGTCGACCGTCGACCGTTTTTTTTTTTTTTT-3'.  cDNAs were
                      cloned into the MluI/SalI sites of a modified pBluescribe
                      vector using commercial linkers (NEB).  Average insert
                      size:  1.0 kb."
BASE COUNT       17 a     14 c     16 g     13 t
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 9;  Length 60;
  Best Local Similarity   83.3%;  Pred. No. 2.2e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       3 CAGGGCACTCATGGCTAT 20
           |||| | |||||||||  |
Db      26 CAGGTCCCTCATGGCTGT 9


RESULT 14
AZ592652
LOCUS       AZ592652                 62 bp    DNA     linear   GSS 13-DEC-2000
DEFINITION  1M0403A13R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0403A13 R, DNA sequence.
ACCESSION   AZ592652
VERSION     AZ592652.1  GI:11714842
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 62)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0403  row: A  column: 13
            Seq primer: CACACAGGAAACAGCTATGACC
            Class: plasmid ends
            High quality sequence stop: 62.
FEATURES              Location/Qualifiers
     source           1. .62
                      /organism="Mus musculus"
                      /strain="C57BL/6J"
                      /db_xref="taxon:10090"
                      /clone="UUGC1M0403A13"
                      /clone_lib="Mouse 10kb plasmid UUGC1M library"
                      /sex="Male"
                      /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                      /note="Vector: PWD42nv; Purified genomic DNA from M.
                      musculus C57BL/6J (male) was obtained  from the Jackson
                      Laboratory Mouse DNA Resource
                      (http://www.jax.org/resources/documents/dnares/). The DNA
                      was hydrodynamically sheared by repeated passage through a
                      0.005 inch orifice at constant velocity. The sheared DNA
                      was blunt end-repaired with T4 DNA polymerase and T4
                      polynucleotide kinase. Adaptor oligonucleotides were
                      ligated to the blunt ends in high  molar excess. The
                      adaptored DNA was purified and size-selected for a 9.5 to
                      10.5 kb range using preparative agarose gel
                      electrophoresis. Vector DNA was prepared from a derivative
                      of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                      inducible derivative of plasmid R1. The vector was ligated
                      with adaptors complementary to the insert adaptors and
                      purified. The sheared, adaptored mouse DNA was annealed to
                      adaptored vector DNA, and transformed into
                      chemically-competent E. coli XL10-Gold (Stratagene) cells
                      and selected for ampicillin resistance."
BASE COUNT       14 a     14 c     20 g     14 t
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 17;  Length 62;
  Best Local Similarity   83.3%;  Pred. No. 2.2e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 CCAGGGCACTCATGGCTA 19
           ||||| |||||  |||||
Db      15 CCAGGCCACTCTCGGCTA 32


RESULT 15
AU260179
LOCUS       AU260179                 75 bp    mRNA    linear   EST 25-APR-2002
DEFINITION  AU260179 3'-directed mouse cDNA library Mus musculus cDNA clone
            BED0016309 3', mRNA sequence.
ACCESSION   AU260179
VERSION     AU260179.1  GI:20327417
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 75)
  AUTHORS   Kato,K. and Matoba,R.
  TITLE     Generation of expressed sequence tags from mouse brain
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Kikuya Kato
            Graduate School of Biological Sciences
            Nara Institute of Science and Technology
            8916-5 Takayama, Ikoma, Nara 630-0101, Japan
            Tel: 81-743-72-5581
            Fax: 81-743-72-5589
            Email: kkato@bs.aist-nara.ac.jp,
            URL:http://love2.aist-nara.ac.jp/BED/index.html.
FEATURES              Location/Qualifiers
     source           1. .75
                      /organism="Mus musculus"
                      /db_xref="taxon:10090"
                      /clone="BED0016309"
                      /clone_lib="3'-directed mouse cDNA library"
                      /tissue_type="brain"
                      /note="Vector: pGEM-T-easy"
BASE COUNT       23 a     18 c     18 g     16 t
ORIGIN
```

Query Match 66.0%; Score 13.2; DB 9; Length 75;
Best Local Similarity 83.3%; Pred. No. 2.4e+04;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy       2 CCAGGGCACTCATGGCTA 19
           ||||| || ||| |||||
Db      54 CCAGGACAGTCAAGGCTA 71
```

Search completed: November 26, 2002, 04:09:37
Job time : 763.8 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 3, 2002, 18:14:22 ; Search time 351.3 Seconds
(without alignments)
1656.863 Million cell updates/sec

Title: US-09-296-264-21
Perfect score: 20
Sequence: 1 gctgagaaaccttcttttgc 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 995600

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
c 1	16.4	82.0	68	6	I25979	I25979 Sequence 11
c 2	15.2	76.0	68	6	I25981	I25981 Sequence 13
c 3	14.6	73.0	93	6	I25978	I25978 Sequence 10
4	14.4	72.0	80	3	DRORS1584	M30167 D.melanogas
5	14.4	72.0	80	3	DRORS1586	M31539 D.melanogas
6	14.2	71.0	20	6	AR072797	AR072797 Sequence
7	14.2	71.0	26	6	AR146081	AR146081 Sequence
8	14	70.0	20	6	I25975	I25975 Sequence 6
9	13.8	69.0	89	8	AF374435	AF374435 Physalis
10	13.8	69.0	89	8	AF374442	AF374442 Physalis
11	13.6	68.0	32	6	A28097	A28097 Human GABAa
12	13.6	68.0	32	6	A39292	A39292 Sequence 17
13	13.6	68.0	32	6	I88873	I88873 Sequence 7
14	13.6	68.0	51	6	AX159778	AX159778 Sequence
15	13.6	68.0	60	9	HUMSAU3A23	D49607 Human alpho
16	13.6	68.0	60	9	HUMSAU3A24	D49608 Human alpho
17	13.6	68.0	60	9	HUMSAU3A27	D49611 Human alpho
18	13.6	68.0	60	9	HUMSAU3A55	D49639 Human alpho
19	13.6	68.0	60	9	HUMSAU3A56	D49640 Human alpho
20	13.6	68.0	60	9	HUMSAU3A67	D49651 Human alpho
c 21	13.6	68.0	68	6	I25980	I25980 Sequence 12
c 22	13.2	66.0	51	6	AX117621	AX117621 Sequence
23	13.2	66.0	72	6	AX350332	AX350332 Sequence
24	13.2	66.0	75	6	AX350330	AX350330 Sequence
c 25	13.2	66.0	81	6	AX107305	AX107305 Sequence
c 26	13.2	66.0	93	11	HUMSWX439	L24655 Human chrom
c 27	12.8	64.0	22	6	A34694	A34694 Synthetic p
c 28	12.8	64.0	29	6	AR050911	AR050911 Sequence
29	12.8	64.0	30	6	AX184118	AX184118 Sequence
30	12.8	64.0	36	6	AX464538	AX464538 Sequence
31	12.8	64.0	48	6	AX300417	AX300417 Sequence
c 32	12.8	64.0	49	6	AR032506	AR032506 Sequence
c 33	12.8	64.0	49	6	AR209170	AR209170 Sequence
c 34	12.8	64.0	49	6	I29246	I29246 Sequence 11
c 35	12.8	64.0	49	6	I90920	I90920 Sequence 11
c 36	12.8	64.0	51	6	A01803	A01803 Nucleotide
c 37	12.8	64.0	51	6	A04006	A04006 Nucleotide
38	12.8	64.0	51	6	AX157282	AX157282 Sequence
39	12.8	64.0	51	6	AX162483	AX162483 Sequence
40	12.8	64.0	54	10	MMDEND54	Y11223 M.musculus
41	12.8	64.0	66	14	SV4EV118	K01104 African gre
42	12.8	64.0	100	10	RNU12520	U12520 Rattus norv
c 43	12.6	63.0	20	6	AR158617	AR158617 Sequence
c 44	12.6	63.0	20	6	AR158618	AR158618 Sequence
45	12.6	63.0	25	6	AX010418	AX010418 Sequence

ALIGNMENTS

RESULT 1
I25979/c
LOCUS I25979 68 bp DNA linear PAT.07-OCT-1996
DEFINITION Sequence 11 from patent US 5554592.
ACCESSION I25979
VERSION I25979.1 GI:1605849
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 68)
AUTHORS Quistad,G.B. and Leisy,D.J.
TITLE Insecticidal toxins from the parastic wasp, bracon hebetor
JOURNAL Patent: US 5554592-A 11 10-SEP-1996;
FEATURES Location/Qualifiers

```
     source          1. .68
                     /organism="unknown"
BASE COUNT       22 a      7 c     11 g     28 t
ORIGIN

  Query Match            82.0%;  Score 16.4;  DB 6;  Length 68;
  Best Local Similarity   94.4%;  Pred. No. 2.2e+03;
  Matches   17;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       3 TGAGAAACCTTCTTTTGC 20
           ||| ||||||||||||||
Db      65 TGAAAAACCTTCTTTTGC 48


RESULT 2
I25981/c
LOCUS       I25981                    68 bp    DNA     linear   PAT 07-OCT-1996
DEFINITION  Sequence 13 from patent US 5554592.
ACCESSION   I25981
VERSION     I25981.1  GI:1605851
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 68)
  AUTHORS   Quistad,G.B. and Leisy,D.J.
  TITLE     Insecticidal toxins from the parastic wasp, bracon hebetor
  JOURNAL   Patent: US 5554592-A 13 10-SEP-1996;
FEATURES             Location/Qualifiers
     source          1. .68
                     /organism="unknown"
BASE COUNT       23 a      8 c     11 g     26 t
ORIGIN

  Query Match            76.0%;  Score 15.2;  DB 6;  Length 68;
  Best Local Similarity   85.0%;  Pred. No. 8.9e+03;
  Matches   17;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 GCTGAGAAACCTTCTTTTGC 20
           | ||| | ||||||||||||
Db      67 GTTGAAATACCTTCTTTTGC 48


RESULT 3
I25978/c
LOCUS       I25978                    93 bp    DNA     linear   PAT 07-OCT-1996
DEFINITION  Sequence 10 from patent US 5554592.
ACCESSION   I25978
VERSION     I25978.1  GI:1605848
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 93)
  AUTHORS   Quistad,G.B. and Leisy,D.J.
  TITLE     Insecticidal toxins from the parastic wasp, bracon hebetor
  JOURNAL   Patent: US 5554592-A 10 10-SEP-1996;
FEATURES             Location/Qualifiers
     source          1. .93
                     /organism="unknown"
BASE COUNT       21 a     10 c      8 g     17 t     37 others
ORIGIN

  Query Match            73.0%;  Score 14.6;  DB 6;  Length 93;
  Best Local Similarity   55.0%;  Pred. No. 1.7e+04;
  Matches   11;  Conservative    7;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 GCTGAGAAACCTTCTTTTGC 20
           ::||: | ||:||:||:||:
Db      70 RYTGRAANACYTTYTTYTGY 51


RESULT 4
```

```
DRORS1584
LOCUS       DRORS1584                 80 bp    DNA     linear   INV 14-SEP-1995
DEFINITION  D.melanogaster nested repetitive sequences F and G, segment 4 of 6.
ACCESSION   M30167
VERSION     M30167.1  GI:158346
KEYWORDS    repeat region; transposon.
SEGMENT     4 of 6
SOURCE      Drosophila melanogaster (clone: pDmI 158.) DNA.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 80)
  AUTHORS   Di Nocera,P.P. and Dawid,I.B.
  TITLE     Interdigitated arrangement of two oligo(A)-terminated DNA sequences
            in Drosophila
  JOURNAL   Nucleic Acids Res. 11 (16), 5475-5482 (1983)
  MEDLINE   83299229
   PUBMED   6310501
FEATURES             Location/Qualifiers
     source          1. .80
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone="pDmI 158."
BASE COUNT       45 a      9 c      5 g     21 t
ORIGIN      About 10 kb after segment 3.

  Query Match            72.0%;  Score 14.4;  DB 3;  Length 80;
  Best Local Similarity   93.8%;  Pred. No. 2.2e+04;
  Matches   15;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       5 AGAAACCTTCTTTTGC 20
           ||||||||||||||| |
Db      62 AGAAACCTTCTTTTAC 77


RESULT 5
DRORS1586
LOCUS       DRORS1586                 80 bp    DNA     linear   INV 14-SEP-1995
DEFINITION  D.melanogaster nested repetitive sequences F and G, segment 6 of 6.
ACCESSION   M31539
VERSION     M31539.1  GI:158348
KEYWORDS    repeat region; transposon.
SEGMENT     6 of 6
SOURCE      Drosophila melanogaster (clone: pDmI 158.) DNA.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 80)
  AUTHORS   Di Nocera,P.P. and Dawid,I.B.
  TITLE     Interdigitated arrangement of two oligo(A)-terminated DNA sequences
            in Drosophila
  JOURNAL   Nucleic Acids Res. 11 (16), 5475-5482 (1983)
  MEDLINE   83299229
   PUBMED   6310501
FEATURES             Location/Qualifiers
     source          1. .80
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone="pDmI 158."
BASE COUNT       45 a      9 c      5 g     21 t
ORIGIN      About 4 kb after segment 5.

  Query Match            72.0%;  Score 14.4;  DB 3;  Length 80;
  Best Local Similarity   93.8%;  Pred. No. 2.2e+04;
  Matches   15;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       5 AGAAACCTTCTTTTGC 20
           ||||||||||||||| |
Db      62 AGAAACCTTCTTTTAC 77
```

```
RESULT 6
AR072797
LOCUS       AR072797                  20 bp    DNA     linear   PAT 28-AUG-2000
DEFINITION  Sequence 3 from patent US 5948650.
ACCESSION   AR072797
VERSION     AR072797.1  GI:9999561
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 20)
  AUTHORS   Araki,S. and Tsuchiya,Y.
  TITLE     Genetic variety identifying method in hops
  JOURNAL   Patent: US 5948650-A 3 07-SEP-1999;
FEATURES             Location/Qualifiers
     source          1. .20
                     /organism="unknown"
BASE COUNT        4 a      4 c      6 g      6 t
ORIGIN

  Query Match             71.0%;  Score 14.2;  DB 6;  Length 20;
  Best Local Similarity   84.2%;  Pred. No. 3.2e+04;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 GCTGAGAAACCTTCTTTTG 19
          |||||| || ||||||| |
Db      2 GCTGAGCAAGCTTCTTTGG 20


RESULT 7
AR146081
LOCUS       AR146081                  26 bp    DNA     linear   PAT 08-AUG-2001
DEFINITION  Sequence 30 from patent US 6218154.
ACCESSION   AR146081
VERSION     AR146081.1  GI:15109270
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 26)
  AUTHORS   Romano,J.W., Shurtliff,R. and Williams,K.G.
  TITLE     Isothermal transcription based assay for the detection and
            quantification of chemokines rantes, MIP-1.alpha. and MIP-1.beta
  JOURNAL   Patent: US 6218154-A 30 17-APR-2001;
FEATURES             Location/Qualifiers
     source          1. .26
                     /organism="unknown"
BASE COUNT        7 a      9 c      2 g      8 t
ORIGIN

  Query Match             71.0%;  Score 14.2;  DB 6;  Length 26;
  Best Local Similarity   84.2%;  Pred. No. 3.1e+04;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 CTGAGAAACCTTCTTTTGC 20
          |||||||| | |||||| |
Db      7 CTGAGAAAACCTCTTTTCC 25


RESULT 8
I25975
LOCUS       I25975                    20 bp    DNA     linear   PAT 07-OCT-1996
DEFINITION  Sequence 6 from patent US 5554592.
ACCESSION   I25975
VERSION     I25975.1  GI:1605845
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 20)
  AUTHORS   Quistad,G.B. and Leisy,D.J.
  TITLE     Insecticidal toxins from the parastic wasp, bracon hebetor
  JOURNAL   Patent: US 5554592-A 6 10-SEP-1996;
FEATURES             Location/Qualifiers
     source          1. .20
                     /organism="unknown"
BASE COUNT        4 a      1 c      2 g      6 t      7 others
ORIGIN

  Query Match             70.0%;  Score 14;  DB 6;  Length 20;
  Best Local Similarity   57.9%;  Pred. No. 4e+04;
  Matches   11;  Conservative    6;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 GCTGAGAAACCTTCTTTTG 19
          ::||: | ||:||:||:||
Db      2 RYTGRAANACYTTYTTYTG 20


RESULT 9
AF374435
LOCUS       AF374435                  89 bp    DNA     linear   PLN 02-MAY-2002
DEFINITION  Physalis longifolia self-incompatibility (S) gene, S-5 allele,
            intron.
ACCESSION   AF374435
VERSION     AF374435.1  GI:20385656
KEYWORDS    .
SOURCE      Physalis longifolia.
  ORGANISM  Physalis longifolia
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Asteridae; euasterids I; Solanales; Solanaceae; Physalis.
REFERENCE   1  (bases 1 to 89)
  AUTHORS   Lu,Y.
  TITLE     Selection and functional constraint at the self-incompatibility
            locus revealed by patterns of sequence diversity in Physalis
            longifolia (Solanaceae)
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 89)
  AUTHORS   Lu,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-APR-2001) Biology, Duke University, Box 90338,
            Durham, NC 27708, USA
FEATURES             Location/Qualifiers
     source          1. .89
                     /organism="Physalis longifolia"
                     /db_xref="taxon:161495"
                     /tissue_type="leaf"
     gene            <1. .>89
                     /gene="S"
                     /note="self-incompatibility"
                     /allele="S-5"
     intron          1. .89
                     /gene="S"
BASE COUNT       25 a     16 c     11 g     37 t
ORIGIN

  Query Match             69.0%;  Score 13.8;  DB 8;  Length 89;
  Best Local Similarity   88.2%;  Pred. No. 4.4e+04;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      2 CTGAGAAACCTTCTTTT 18
          |||| ||||||| ||||
Db     59 CTGAAAAACCTTATTTT 75


RESULT 10
AF374442
LOCUS       AF374442                  89 bp    DNA     linear   PLN 02-MAY-2002
DEFINITION  Physalis longifolia self-incompatibility (S) gene, S-13 allele,
            intron.
ACCESSION   AF374442
VERSION     AF374442.1  GI:20385663
KEYWORDS    .
SOURCE      Physalis longifolia.
  ORGANISM  Physalis longifolia
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
```

```
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Asteridae; euasterids I; Solanales; Solanaceae; Physalis.
REFERENCE   1  (bases 1 to 89)
  AUTHORS   Lu,Y.
  TITLE     Selection and functional constraint at the self-incompatibility
            locus revealed by patterns of sequence diversity in Physalis
            longifolia (Solanaceae)
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 89)
  AUTHORS   Lu,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-APR-2001) Biology, Duke University, Box 90338,
            Durham, NC 27708, USA
FEATURES             Location/Qualifiers
     source          1. .89
                     /organism="Physalis longifolia"
                     /db_xref="taxon:161495"
                     /tissue_type="leaf"
     gene            <1. .>89
                     /gene="S"
                     /note="self-incompatibility"
                     /allele="S-13"
     intron          1. .89
                     /gene="S"
BASE COUNT       23 a     19 c     11 g     36 t
ORIGIN

  Query Match             69.0%;  Score 13.8;  DB 8;  Length 89;
  Best Local Similarity   88.2%;  Pred. No. 4.4e+04;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      2 CTGAGAAACCTTCTTTT 18
          |||| ||||||| ||||
Db     59 CTGAAAAACCTTATTTT 75


RESULT 11
A28097
LOCUS       A28097                     32 bp    DNA     linear   PAT 11-OCT-1995
DEFINITION  Human GABAa beta-2 subunit PCR primer.
ACCESSION   A28097
VERSION     A28097.1  GI:1248595
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
FEATURES             Location/Qualifiers
     source          1. .32
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
BASE COUNT       13 a      6 c      7 g      6 t
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 6;  Length 32;
  Best Local Similarity   80.0%;  Pred. No. 6.1e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 GCTGAGAAACCTTCTTTTGC 20
          |||||||||  ||  ||  |||
Db     13 GCTGAGAAAGCTGCTAATGC 32


RESULT 12
A39292
LOCUS       A39292                     32 bp    DNA     linear   PAT 05-MAR-1997
DEFINITION  Sequence 17 from Patent WO9413799.
ACCESSION   A39292
VERSION     A39292.1  GI:2295653
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 32)
  AUTHORS   Hadingham,K.L. and Whiting,P.J.
  TITLE     STABLY TRANSFECTED CELL LINES EXPRESSING GABA-A RECEPTORS
  JOURNAL   Patent: WO 9413799-A 17 23-JUN-1994;
            MERCK SHARP & DOHME (GB)
COMMENT     Other publication CA 2151236 940623
            Other publication AU 5655494 940704
            Other publication JP 8504330T 960514.
FEATURES             Location/Qualifiers
     source          1. .32
                     /organism="unidentified"
                     /db_xref="taxon:32644"
BASE COUNT       13 a      6 c      7 g      6 t
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 6;  Length 32;
  Best Local Similarity   80.0%;  Pred. No. 6.1e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 GCTGAGAAACCTTCTTTTGC 20
          |||||||||  ||  ||  |||
Db     13 GCTGAGAAAGCTGCTAATGC 32


RESULT 13
I88873
LOCUS       I88873                     32 bp    DNA     linear   PAT 10-AUG-1998
DEFINITION  Sequence 7 from patent US 5719057.
ACCESSION   I88873
VERSION     I88873.1  GI:3408813
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 32)
  AUTHORS   Hadingham,K.Louise., le Bourdelles,B., Whiting,P.John. and
            Wingrove,P.Baxter.
  TITLE     Stably human transfected rodent fibroblast cell line expressing
            human GABA-A recepotors, and cloned human GABA-A receptor subunit
            CDNA sequences
  JOURNAL   Patent: US 5719057-A 7 17-FEB-1998;
FEATURES             Location/Qualifiers
     source          1. .32
                     /organism="unknown"
BASE COUNT       13 a      6 c      7 g      6 t
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 6;  Length 32;
  Best Local Similarity   80.0%;  Pred. No. 6.1e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 GCTGAGAAACCTTCTTTTGC 20
          |||||||||  ||  ||  |||
Db     13 GCTGAGAAAGCTGCTAATGC 32


RESULT 14
AX159778
LOCUS       AX159778                   51 bp    DNA     linear   PAT 22-JUN-2001
DEFINITION  Sequence 3106 from Patent WO0140521.
ACCESSION   AX159778
VERSION     AX159778.1  GI:14541109
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 51)
  AUTHORS   Shimkets,R.A. and Leach,M.
  TITLE     Nucleic acids containing single nucleotide polymorphisms and
            methods of use thereof
  JOURNAL   Patent: WO 0140521-A 3106 07-JUN-2001;
            Curagen Corporation (US)
FEATURES             Location/Qualifiers
```

```
     source          1. .51
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     misc_feature    26
                     /note="2 of 2 allelic variants (3105 is other entry)
                     Accession number cg42913480"
BASE COUNT       13 a      7 c      8 g     23 t
ORIGIN

  Query Match            68.0%;  Score 13.6;  DB 6;  Length 51;
  Best Local Similarity  80.0%;  Pred. No. 5.8e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 GCTGAGAAACCTTCTTTTGC 20
           ||| |||||  ||||| |||
Db      13 GCTAAGAAAAATTCTTCTGC 32


RESULT 15
HUMSAU3A23
LOCUS       HUMSAU3A23               60 bp    DNA     linear   PRI 10-FEB-1999
DEFINITION  Human alphoid Sau3A repeat family gene.
ACCESSION   D49607
VERSION     D49607.1  GI:725405
KEYWORDS    alphoid Sau3A repetitive family DNA; recombination site.
SOURCE      Homo sapines (strain HeLa cell) DNA, clone III-15.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 60)
  AUTHORS   Ohki,R., Oishi,M. and Kiyama,R.
  TITLE     Preference of the recombination sites involved in the formation of
            extrachromosomal copies of the human alphoid Sau3A repeat family
  JOURNAL   Nucleic Acids Res. 23 (24), 4971-4977 (1995)
  MEDLINE   96145252
REFERENCE   2  (bases 1 to 60)
  AUTHORS   Ohki,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-MAR-1995) Rieko Ohki, University of Tokyo, Inst.
            Molecular & Cellular Biosciences; Yayoi 1-1-1, Bunkyo-ku, Tokyo
            113, Japan (Tel:03-3812-2111(ex.7835), Fax:03-3818-9437)
COMMENT     Submitted (07-Mar-1995) to DDBJ by:
            Rieko Ohki
            Inst. Mol. Cell. Biosci.
            Univ. Tokyo
            Yayoi 1-1-1
            Bunkyo-ku, Tokyo 113
            Japan
            Phone: 03-3812-2111 x7835
            Fax:   03-3818-9437.
FEATURES             Location/Qualifiers
     source          1. .60
                     /organism="Homo sapiens"
                     /strain="HeLa cell"
                     /db_xref="taxon:9606"
                     /note="alphoid Sau3A repetitive family DNA-recombination
                     junction-recombination site"
BASE COUNT       16 a     14 c     12 g     18 t
ORIGIN

  Query Match            68.0%;  Score 13.6;  DB 9;  Length 60;
  Best Local Similarity  80.0%;  Pred. No. 5.7e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 GCTGAGAAACCTTCTTTTGC 20
           |||  |||||| |||||| |
Db      12 GCTCTGAAACCCTCTTTTTC 31


Search completed: December  3, 2002, 22:23:01
Job time : 357.3 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 06:29:45 ; Search time 94.1 Seconds
(without alignments)
478.639 Million cell updates/sec

Title: US-09-296-264-21
Perfect score: 20
Sequence: 1 gctgagaaaccttcttttgc 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 2390332

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		20	100.0	20	21	AAZ31451	Human neuropilin m
c 2		16.4	82.0	68	14	AAQ48443	Brh-I N-terminal r
c 3		16.4	82.0	68	22	AAS05989	Nucleotide sequenc
4		15.2	76.0	60	24	ABN32273	Human spliced tran
c 5		15.2	76.0	68	14	AAQ48445	Brh-I N-terminal r
c 6		15.2	76.0	68	22	AAS05991	Nucleotide sequenc
c 7		14.6	73.0	93	14	AAQ48442	Brh-I N-terminal r
c 8		14.6	73.0	93	22	AAS05988	Degenerate DNA seq
9		14.4	72.0	41	22	AAH48165	Fumarate lyase 13
c 10		14.4	72.0	60	24	ABN42595	Human spliced tran
11		14.2	71.0	20	18	AAT66259	Primer 3 for hop g
12		14.2	71.0	26	20	AAX88476	Human MIP-1 beta p
13		14.2	71.0	29	21	AAD00984	Primer #2 for AtEM
14		14	70.0	20	14	AAQ48439	Brh-I N-terminal p
15		14	70.0	20	17	AAT45676	Bracon hebetor Brh
16		14	70.0	20	22	AAS05985	Degenerate PCR pri
c 17		14	70.0	85	21	AAC07306	Human secreted pro
18		13.6	68.0	32	14	AAQ33123	Beta-2 subunit pri
19		13.6	68.0	32	15	AAQ69131	Human GABA recepto
20		13.6	68.0	51	22	AAI76165	Human silent SNP c
21		13.6	68.0	65	24	ABN30451	Rat spliced transc
c 22		13.6	68.0	68	14	AAQ48444	Brh-I N-terminal r
c 23		13.6	68.0	68	22	AAS05990	Nucleotide sequenc
c 24		13.4	67.0	60	24	ABN38158	Human spliced tran
25		13.4	67.0	65	24	ABN52110	Mouse spliced tran
26		13.2	66.0	30	19	AAV27061	Aspergillus oryzae
27		13.2	66.0	32	22	AAH41146	PCR primer #1. Sy
c 28		13.2	66.0	41	24	ABN85540	Human GARP protein
c 29		13.2	66.0	44	24	ABN71778	Streptococcus agal
c 30		13.2	66.0	51	22	AAH39948	Human SNP flanking
31		13.2	66.0	60	24	ABN59287	Human spliced tran
32		13.2	66.0	60	24	ABN59411	Human spliced tran
c 33		13.2	66.0	65	24	ABN54577	Mouse spliced tran
34		13.2	66.0	72	24	ABA91347	Sheep growth diffe
35		13.2	66.0	75	24	ABA91346	Sheep growth diffe
c 36		13.2	66.0	81	22	AAH49929	Bacterial 23S/5S R
37		13.2	66.0	81	24	ABK52384	Fucose-specific le
c 38		13.2	66.0	90	22	AAK18644	Human brain expres
c 39		13.2	66.0	90	22	AAK44572	Human bone marrow
c 40		13.2	66.0	90	22	AAI50555	Probe #19241 used
c 41		13.2	66.0	90	24	ABS18809	Human genome-deriv
c 42		12.8	64.0	20	20	AAX32818	Primer specific fo
c 43		12.8	64.0	24	13	AAQ25465	Purine rich RSV ta
c 44		12.8	64.0	24	24	ABL41526	Primer #1 relating
45		12.8	64.0	29	20	AAZ12379	PCR primer used to

ALIGNMENTS

RESULT 1
AAZ31451
ID AAZ31451 standard; DNA; 20 BP.
XX
AC AAZ31451;
XX
DT 07-FEB-2000 (first entry)
XX
DE Human neuropilin mRNA specific antisense oligo GTI3622.
XX
KW Neuropilin; human; growth; metastasis; tumor; neovascularisation;
KW cancer; papilloma; diabetic retinopathy; antisense; ss.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9955855-A2.
XX
PD 04-NOV-1999.
XX
PF 23-APR-1999; 99WO-CA00324.
XX
PR 23-APR-1998; 98US-0082791.
XX
PA (GENE-) GENESENSE TECHNOLOGIES INC.
XX
PI Wright JA, Young AH, Lee YS;
XX
DR WPI; 2000-023357/02.
XX
PT Antisense oligonucleotides that inhibit neuropilin expression, useful
PT for treating cancer -

```
XX
PS   Claim 4; Page 16; 57pp; English.
XX
CC   Sequences AAZ31431-460 represent antisense oligonucleotides which
CC   inhibit human neuropilin expression. The antisense oligonucleotides can
CC   be used to inhibit the growth or metastasis of a mammalian tumor and
CC   inhibit neovascularisation. The oligonucleotides may be used to treat
CC   various forms of cancers or tumors, such as sarcomas, melanomas,
CC   adenomas, carcinomas of solid tissue, hypoxic tumors, squamous cell
CC   carcinomas of the mouth, throat, larynx and lung, genitourinary cancers
CC   such as cervical and bladder cancer, hematopoietic cancers, colon cancer,
CC   breast cancer, pancreatic cancer, renal cancer, brain cancer, skin
CC   cancer, liver cancer, head and neck cancers, and nervous system cancers,
CC   as well as benign lesions such as papillomas. The methods may be used to
CC   treat neovascularisation disorders such as diabetic retinopathy, and
CC   retinopathy of prematurity and age related macular degeneration.
XX
SQ   Sequence 20 BP; 4 A; 5 C; 4 G; 7 T; 0 other;

  Query Match             100.0%;  Score 20;  DB 21;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 2.2;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 GCTGAGAAACCTTCTTTTGC 20
           ||||||||||||||||||||
Db       1 GCTGAGAAACCTTCTTTTGC 20


RESULT 2
AAQ48443/c
ID   AAQ48443 standard; DNA; 68 BP.
XX
AC   AAQ48443;
XX
DT   23-MAR-1994  (first entry)
XX
DE   Brh-I N-terminal reverse translation sequence #1.
XX
KW   Wasp; insect; toxicity; ectoparasite; Habrobracon hebetor; venom; PCR;
KW   Microbracon hebetor; paralysis; Lepidoptera; polymerase chain reaction;
KW   primer; amplify; Bracon hebetor; ss.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   primer_bind     1..20
FT                   /*tag=  a
FT   primer_bind     complement (49..68)
FT                   /*tag=  b
XX
PN   WO9318145-A.
XX
PD   16-SEP-1993.
XX
PF   25-FEB-1993;    93WO-EP00431.
XX
PR   04-MAR-1992;    92US-0847570.
PR   10-JUN-1992;    92US-0897192.
XX
PA   (SANO ) SANDOZ LTD.
PA   (SANO ) SANDOZ PATENT GMBH.
PA   (SANO ) SANDOZ-ERFINDUNGEN VERW GES MBH.
XX
PI   Leisy DJ,  Quistad GB;
XX
DR   WPI; 1993-303459/38.
XX
PT   Polypeptide(s) from Bracon hebetor venom - useful as insecticides,
PT   when cloned into baculoviruses
XX
PS   Example 5; Page 13; 43pp; English.
XX
CC   The sequences given in AAQ48443-45 represent nucleotide sequences
CC   derived from the reverse translation product of the Bracon hebetor
CC   toxin, Brh-I, N-terminal amino acid sequence.  These sequences were
CC   used in the production of primers.  These primers were used in the
CC   production of Brh-I cDNA.  The amplification product encodes a wasp
CC   polypeptide which exhibits insect toxicity and has a molecular weight
CC   >70 kD.  This sequence was derived from the ectoparasitic wasp also
CC   known as Habrobracon hebetor and Microbracon hebetor.  The protein
CC   encoded by the amplification product is a venom which is paralytic
CC   to Lepidopterans, and may therefore be used for controlling insects,
CC   esp. Lepidoptera.  This protein is quite labile under cetain isolation
CC   conditions, it is particularly unstable at low pH.
XX
SQ   Sequence 68 BP; 22 A; 7 C; 11 G; 28 T; 0 other;

  Query Match             82.0%;  Score 16.4;  DB 14;  Length 68;
  Best Local Similarity   94.4%;  Pred. No. 1.5e+02;
  Matches   17;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       3 TGAGAAACCTTCTTTTGC 20
           ||| ||||||||||||||
Db      65 TGAAAAACCTTCTTTTGC 48


RESULT 3
AAS05989/c
ID   AAS05989 standard; DNA; 68 BP.
XX
AC   AAS05989;
XX
DT   23-OCT-2001  (first entry)
XX
DE   Nucleotide sequence of B. hebetor Brh-I cloned PCR fragment #1.
XX
KW   Insecticidal toxin; Brh-I; parasitic wasp venom; insect paralysis;
KW   larvae; insecticide; Lepidoptera; Spodoptera; ds.
XX
OS   Bracon hebetor.
XX
FH   Key             Location/Qualifiers
FT   primer_bind     complement (1..20)
FT                   /*tag=  a
FT   primer_bind     52..68
FT                   /*tag=  b
XX
PN   US6251862-B1.
XX
PD   26-JUN-2001.
XX
PF   31-MAY-1996;    96US-0655782.
XX
PR   05-FEB-1993;    93US-0013890.
PR   10-AUG-1994;    94US-0288408.
PR   04-MAR-1992;    92US-0847570.
PR   10-JUN-1992;    92US-0897192.
XX
PA   (SYGN ) SYNGENTA PARTICIPATIONS AG.
XX
PI   Quistad GB,  Leisy DJ;
XX
DR   WPI; 2001-424486/45.
XX
PT   Bracon hebetor nucleic acid molecule encoding a toxic Brh-I
PT   polypeptide, useful for controlling insects of order Lepidoptera, and
PT   of genus Spodoptera -
XX
PS   Example 5; Column 9-10; 23pp; English.
XX
CC   The present sequence represents the nucleotide sequence of Bracon
CC   hebetor Brh-I cloned PCR fragment #1. Brh-I (AAU03582) is a
CC   novel insecticidal toxin which is isolated from the venom of the
CC   parasitic wasp B. hebetor. Brh-I and 4 other toxins are purified from
CC   B. hebetor venom glands using anion-exchange chromatography followed
CC   by reverse-phase HPLC and sequencing to identify the various toxins.
```

CC Partial sequences for Brh-I (AAU03579), Brh-V (AAU03580) and Brh-III
CC (AAU03581) obtained by this method are given. Brh-I and the 4 other
CC Brh toxins induce paralysis and/or death of insect larvae. The
CC polynucleotide encoding Brh-I toxin can be cloned into a baculovirus
CC and is useful for recombinantly expressing Brh-I. Brh-I is useful as
CC an insecticide against insects of the order Lepidoptera e.g. Heliothis
CC virescens, Autographa californica and insects of the genus Spodoptera.
CC The toxic polypeptides are toxic, i.e. paralytic, and/or lethal at very
CC low concentrations.
XX
SQ Sequence 68 BP; 22 A; 7 C; 11 G; 28 T; 0 other;

Query Match 82.0%; Score 16.4; DB 22; Length 68;
Best Local Similarity 94.4%; Pred. No. 1.5e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 TGAGAAACCTTCTTTTGC 20
||| ||||||||||||||
Db 65 TGAAAAACCTTCTTTTGC 48

RESULT 4
ABN32273
ID ABN32273 standard; DNA; 60 BP.
XX
AC ABN32273;
XX
DT 15-JUL-2002 (first entry)
XX
DE Human spliced transcript detection oligonucleotide SEQ ID NO:5021.
XX
KW Human; mouse; rat; splice transcript; detection; RNA transcript;
KW splice variant; transcriptome; oligonucleotide library; ss.
XX
OS Homo sapiens.
XX
PN WO200210449-A2.
XX
PD 07-FEB-2002.
XX
PF 20-JUL-2001; 2001WO-IB01903.
XX
PR 28-JUL-2000; 2000US-221607P.
PR 02-MAY-2001; 2001US-287724P.
XX
PA (COMP-) COMPUGEN INC.
XX
PI Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
XX
DR WPI; 2002-257383/30.
XX
PT New oligonucleotide libraries comprising oligonucleotides which
PT selectively hybridize to mRNAs transcribed from a transcription unit of
PT a genome, useful for detecting tissue-, pathology-, and
PT developmental-specific genes -
XX
PS Example 1; SEQ ID 5021; 47pp; English.
XX
CC The present invention describes oligonucleotide libraries for detecting
CC messenger RNAs that populate a (sub-)transcriptome, where the
CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC transcription units that populate a genome. The library comprises
CC several oligonucleotides, each capable of hybridising selectively to a
CC set of messenger RNAs transcribed from a given transcription unit of
CC the genome, which encodes one or more messenger RNA splice variants.
CC The oligonucleotide libraries are useful for detecting mRNAs from a
CC biological sample, in expression profiling studies, in qualitatively or
CC quantitatively characterising the corresponding transcriptome, and in
CC detecting RNA transcripts and splice variants of human or animal
CC transcriptomes. The libraries may also be used as specialised mini
CC libraries to detect transcripts of a sub-transcriptome under a
CC particular biological or pathological state, and so allowing the
CC detection of tissue- and pathology-specific genes such as those genes
CC only expressed in specific tissue under a specific pathological
CC condition; to detect developmental specific genes; and to detect RNA
CC transcripts and splice variants of a transcriptome of a patient suffering
CC from a particular disorder. ABN27253 to ABN59589 represent
CC oligonucleotide sequences from rats, humans and mice, which are used in
CC the exemplification of the present invention.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 60 BP; 14 A; 16 C; 14 G; 16 T; 0 other;

Query Match 76.0%; Score 15.2; DB 24; Length 60;
Best Local Similarity 85.0%; Pred. No. 5.7e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GCTGAGAAACCTTCTTTTGC 20
||||||||| |||| ||||
Db 11 GCTGAGAAAGGTTCTCTTGC 30

RESULT 5
AAQ48445/c
ID AAQ48445 standard; DNA; 68 BP.
XX
AC AAQ48445;
XX
DT 23-MAR-1994 (first entry)
XX
DE Brh-I N-terminal reverse translation sequence #3.
XX
KW Wasp; insect; toxicity; ectoparasite; Habrobracon hebetor; venom; PCR;
KW Microbracon hebetor; paralysis; Lepidoptera; polymerase chain reaction;
KW primer; amplify; Bracon hebetor; ss.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT primer_bind 1..20
FT /*tag= a
FT primer_bind complement (49..68)
FT /*tag= b
XX
PN WO9318145-A.
XX
PD 16-SEP-1993.
XX
PF 25-FEB-1993; 93WO-EP00431.
XX
PR 04-MAR-1992; 92US-0847570.
PR 10-JUN-1992; 92US-0897192.
XX
PA (SANO) SANDOZ LTD.
PA (SANO) SANDOZ PATENT GMBH.
PA (SANO) SANDOZ-ERFINDUNGEN VERW GES MBH.
XX
PI Leisy DJ, Quistad GB;
XX
DR WPI; 1993-303459/38.
XX
PT Polypeptide(s) from Bracon hebetor venom - useful as insecticides,
PT when cloned into baculoviruses
XX
PS Example 5; Page 13; 43pp; English.
XX
CC The sequences given in AAQ48443-45 represent nucleotide sequences
CC derived from the reverse translation product of the Bracon hebetor
CC toxin, Brh-I, N-terminal amino acid sequence. These sequences were
CC used in the production of primers. These primers were used in the
CC production of Brh-I cDNA. The amplification product encodes a wasp
CC polypeptide which exhibits insect toxicity and has a molecular weight
CC >70 kD. This sequence was derived from the ectoparasitic wasp also
CC known as Habrobracon hebetor and Microbracon hebetor. The protein

```
CC   encoded by the amplification product is a venom which is paralytic
CC   to Lepidopterans, and may therefore be used for controlling insects,
CC   esp. Lepidoptera.  This protein is quite labile under cetain isolation
CC   conditions, it is particularly unstable at low pH.
XX
SQ   Sequence 68 BP; 23 A; 8 C; 11 G; 26 T; 0 other;

  Query Match             76.0%;  Score 15.2;  DB 14;  Length 68;
  Best Local Similarity   85.0%;  Pred. No. 5.7e+02;
  Matches   17;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 GCTGAGAAACCTTCTTTTGC 20
          | ||| | ||||||||||||
Db     67 GTTGAAATACCTTCTTTTGC 48


RESULT 6
AAS05991/c
ID   AAS05991 standard; DNA; 68 BP.
XX
AC   AAS05991;
XX
DT   23-OCT-2001  (first entry)
XX
DE   Nucleotide sequence of B. hebetor Brh-I cloned PCR fragment #3.
XX
KW   Insecticidal toxin; Brh-I; parasitic wasp venom; insect paralysis;
KW   larvae; insecticide; Lepidoptera; Spodoptera; ds.
XX
OS   Bracon hebetor.
XX
FH   Key             Location/Qualifiers
FT   primer_bind     complement (1..20)
FT                   /*tag=  a
FT   primer_bind     52..68
FT                   /*tag=  b
XX
PN   US6251862-B1.
XX
PD   26-JUN-2001.
XX
PF   31-MAY-1996;   96US-0655782.
XX
PR   05-FEB-1993;   93US-0013890.
PR   10-AUG-1994;   94US-0288408.
PR   04-MAR-1992;   92US-0847570.
PR   10-JUN-1992;   92US-0897192.
XX
PA   (SYGN ) SYNGENTA PARTICIPATIONS AG.
XX
PI   Quistad GB,  Leisy DJ;
XX
DR   WPI; 2001-424486/45.
XX
PT   Bracon hebetor nucleic acid molecule encoding a toxic Brh-I
PT   polypeptide, useful for controlling insects of order Lepidoptera, and
PT   of genus Spodoptera -
XX
PS   Example 5; Column 9-10; 23pp; English.
XX
CC   The present sequence represents the nucleotide sequence of Bracon
CC   hebetor Brh-I cloned PCR fragment #3. Brh-I (AAU03582) is a
CC   novel insecticidal toxin which is isolated from the venom of the
CC   parasitic wasp B. hebetor. Brh-I and 4 other toxins are purified from
CC   B. hebetor venom glands using anion-exchange chromatography followed
CC   by reverse-phase HPLC and sequencing to identify the various toxins.
CC   Partial sequences for Brh-I (AAU03579), Brh-V (AAU03580) and Brh-III
CC   (AAU03581) obtained by this method are given. Brh-I and the 4 other
CC   Brh toxins induce paralysis and/or death of insect larvae. The
CC   polynucleotide encoding Brh-I toxin can be cloned into a baculovirus
CC   and is useful for recombinantly expressing Brh-I. Brh-I is useful as
CC   an insecticide against insects of the order Lepidoptera e.g. Heliothis
CC   virescens, Autographa californica and insects of the genus Spodoptera.
CC   The toxic polypeptides are toxic, i.e. paralytic, and/or lethal at very
CC   low concentrations.
XX
SQ   Sequence 68 BP; 23 A; 8 C; 11 G; 26 T; 0 other;

  Query Match             76.0%;  Score 15.2;  DB 22;  Length 68;
  Best Local Similarity   85.0%;  Pred. No. 5.7e+02;
  Matches   17;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 GCTGAGAAACCTTCTTTTGC 20
          | ||| | ||||||||||||
Db     67 GTTGAAATACCTTCTTTTGC 48


RESULT 7
AAQ48442/c
ID   AAQ48442 standard; DNA; 93 BP.
XX
AC   AAQ48442;
XX
DT   23-MAR-1994  (first entry)
XX
DE   Brh-I N-terminal reverse translation degenerate sequence.
XX
KW   Wasp; insect; toxicity; ectoparasite; Habrobracon hebetor; venom; PCR;
KW   Microbracon hebetor; paralysis; Lepidoptera; polymerase chain reaction;
KW   primer; amplify; Bracon hebetor; ss.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   primer_bind     4..23
FT                   /*tag=  a
FT   primer_bind     complement (52..71)
FT                   /*tag=  b
XX
PN   WO9318145-A.
XX
PD   16-SEP-1993.
XX
PF   25-FEB-1993;   93WO-EP00431.
XX
PR   04-MAR-1992;   92US-0847570.
PR   10-JUN-1992;   92US-0897192.
XX
PA   (SANO ) SANDOZ LTD.
PA   (SANO ) SANDOZ PATENT GMBH.
PA   (SANO ) SANDOZ-ERFINDUNGEN VERW GES MBH.
XX
PI   Leisy DJ,  Quistad GB;
XX
DR   WPI; 1993-303459/38.
XX
PT   Polypeptide(s) from Bracon hebetor venom - useful as insecticides,
PT   when cloned into baculoviruses
XX
PS   Example 5; Page 13; 43pp; English.
XX
CC   This sequence represents the nucleotide sequence derived by reverse
CC   translation of the Bracon hebetor toxin, Brh-I, amino acid sequence.
CC   This sequence, and more specific sequences derived from it (see also
CC   AAQ48443-45), were used in the production of primers.  These primers
CC   were used in the production of Brh-I cDNA.  The amplification product
CC   encodes a wasp polypeptide which exhibits insect toxicity and has a
CC   molecular weight >70 kD.  This sequence was derived from the
CC   ectoparasitic wasp, also known as Habrobracon hebetor and Microbracon
CC   hebetor.  The protein encoded by the amplification product is a venom
CC   which is paralytic to Lepidopterans, and may therefore be used for
CC   controlling insects, esp. Lepidoptera.  This protein is quite labile
CC   under cetain isolation conditions, it is particularly unstable at low
CC   pH.
XX
SQ   Sequence 93 BP; 21 A; 10 C; 8 G; 17 T; 37 other;
```

```
  Query Match             73.0%;  Score 14.6;  DB 14;  Length 93;
  Best Local Similarity   55.0%;  Pred. No. 1.2e+03;
  Matches   11;  Conservative    7;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 GCTGAGAAACCTTCTTTTGC 20
          ::||: | ||:||:||:||:
Db     70 RYTGRAANACYTTYTTYTGY 51


RESULT 8
AAS05988/c
ID   AAS05988 standard; DNA; 93 BP.
XX
AC   AAS05988;
XX
DT   23-OCT-2001  (first entry)
XX
DE   Degenerate DNA sequence encoding N-terminus of B. hebetor mature Brh-I.
XX
KW   Insecticidal toxin; Brh-I; parasitic wasp venom; insect paralysis;
KW   larvae; insecticide; Lepidoptera; Spodoptera; ds.
XX
OS   Bracon hebetor.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..93
FT                   /*tag= a
FT                   /partial
FT                   /product= "Brh-I N-terminus"
FT                   /note= "This sequence lacks both start and stop codons"
FT   primer_bind     complement (4..23)
FT                   /*tag= b
FT   unsure          43..48
FT                   /*tag= c
FT                   /note= "Encodes Leu-Leu"
FT   primer_bind     55..71
FT                   /*tag= d
FT   unsure          70..75
FT                   /*tag= e
FT                   /note= "Encodes Leu-Leu"
XX
PN   US6251862-B1.
XX
PD   26-JUN-2001.
XX
PF   31-MAY-1996;   96US-0655782.
XX
PR   05-FEB-1993;   93US-0013890.
PR   10-AUG-1994;   94US-0288408.
PR   04-MAR-1992;   92US-0847570.
PR   10-JUN-1992;   92US-0897192.
XX
PA   (SYGN ) SYNGENTA PARTICIPATIONS AG.
XX
PI   Quistad GB,  Leisy DJ;
XX
DR   WPI; 2001-424486/45.
DR   P-PSDB; AAU03583.
XX
PT   Bracon hebetor nucleic acid molecule encoding a toxic Brh-I
PT   polypeptide, useful for controlling insects of order Lepidoptera, and
PT   of genus Spodoptera -
XX
PS   Example 5; Column 9-10; 23pp; English.
XX
CC   The present sequence encoding for N-terminal amino acids 2-32 of
CC   Bracon hebetor mature Brh-I toxin is derived by reverse translation
CC   of the N-terminal Brh-I amino acid sequence. Brh-I (AAU03582) is a
CC   novel insecticidal toxin which is isolated from the venom of the
CC   parasitic wasp B. hebetor. Brh-I and 4 other toxins are purified from
CC   B. hebetor venom glands using anion-exchange chromatography followed
CC   by reverse-phase HPLC and sequencing to identify the various toxins.
CC   Partial sequences for Brh-I (AAU03579), Brh-V (AAU03580) and Brh-III
CC   (AAU03581) obtained by this method are given. Brh-I and the 4 other
CC   Brh toxins induce paralysis and/or death of insect larvae. The
CC   polynucleotide encoding Brh-I toxin can be cloned into a baculovirus
CC   and is useful for recombinantly expressing Brh-I. Brh-I is useful as
CC   an insecticide against insects of the order Lepidoptera e.g. Heliothis
CC   virescens, Autographa californica and insects of the genus Spodoptera.
CC   The toxic polypeptides are toxic, i.e. paralytic, and/or lethal at very
CC   low concentrations.
XX
SQ   Sequence 93 BP; 21 A; 10 C; 8 G; 17 T; 37 other;

  Query Match             73.0%;  Score 14.6;  DB 22;  Length 93;
  Best Local Similarity   55.0%;  Pred. No. 1.2e+03;
  Matches   11;  Conservative    7;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 GCTGAGAAACCTTCTTTTGC 20
          ::||: | ||:||:||:||:
Db     70 RYTGRAANACYTTYTTYTGY 51


RESULT 9
AAH48165
ID   AAH48165 standard; DNA; 41 BP.
XX
AC   AAH48165;
XX
DT   19-SEP-2001  (first entry)
XX
DE   Fumarate lyase 13 probe #1.
XX
KW   Fumarate lyase 13; cytostatic; antiviral; immunomodulatory;
KW   antiinflammatory; malignant neoplasm; haemopathy; HIV infection;
KW   immunological disease; inflammatory disease; probe; ss.
XX
OS   Unidentified.
XX
PN   WO200148177-A1.
XX
PD   05-JUL-2001.
XX
PF   18-DEC-2000; 2000WO-CN00636.
XX
PR   24-DEC-1999;   99CN-0125775.
XX
PA   (BIOW-) BIOWINDOW GENE DEV INC SHANGHAI.
XX
PI   Mao Y,  Xie Y;
XX
DR   WPI; 2001-418272/44.
XX
PT   Fumarate lyase 13 polynucleotide and polypeptide, useful in diagnosis
PT   and treatment of malignant neoplasm, hemopathy, HIV infection,
PT   immunological diseases and various inflammatory diseases -
XX
PS   Example 7; Page 14; 36pp; Chinese.
XX
CC   The present invention relates to fumarate lyase 13 and its coding
CC   sequence (see AAH48160 and AAG64251). The lyase and its coding sequence
CC   are useful in the diagnosis and treatment of malignant neoplasm,
CC   haemopathy, HIV infection, immunological diseases and various
CC   inflammatory diseases. The present sequence is a probe, which was
CC   used in an example from the present invention.
XX
SQ   Sequence 41 BP; 14 A; 9 C; 6 G; 12 T; 0 other;

  Query Match             72.0%;  Score 14.4;  DB 22;  Length 41;
  Best Local Similarity   93.8%;  Pred. No. 1.4e+03;
  Matches   15;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      3 TGAGAAACCTTCTTTT 18
          |||||||||||||| ||
Db     19 TGAGAAACCTTCTCTT 34
```

```
RESULT 10
ABN42595/c
ID   ABN42595 standard; DNA; 60 BP.
XX
AC   ABN42595;
XX
DT   15-JUL-2002  (first entry)
XX
DE   Human spliced transcript detection oligonucleotide SEQ ID NO:15343.
XX
KW   Human; mouse; rat; splice transcript; detection; RNA transcript;
KW   splice variant; transcriptome; oligonucleotide library; ss.
XX
OS   Homo sapiens.
XX
PN   WO200210449-A2.
XX
PD   07-FEB-2002.
XX
PF   20-JUL-2001; 2001WO-IB01903.
XX
PR   28-JUL-2000; 2000US-221607P.
PR   02-MAY-2001; 2001US-287724P.
XX
PA   (COMP-) COMPUGEN INC.
XX
PI   Shoshan A,  Wasserman A,  Mintz E,  Mintz L,  Faigler S;
XX
DR   WPI; 2002-257383/30.
XX
PT   New oligonucleotide libraries comprising oligonucleotides which
PT   selectively hybridize to mRNAs transcribed from a transcription unit of
PT   a genome, useful for detecting tissue-, pathology-, and
PT   developmental-specific genes  -
XX
PS   Example 1; SEQ ID 15343; 47pp; English.
XX
CC   The present invention describes oligonucleotide libraries for detecting
CC   messenger RNAs that populate a (sub-)transcriptome, where the
CC   (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC   transcription units that populate a genome. The library comprises
CC   several oligonucleotides, each capable of hybridising selectively to a
CC   set of messenger RNAs transcribed from a given transcription unit of
CC   the genome, which encodes one or more messenger RNA splice variants.
CC   The oligonucleotide libraries are useful for detecting mRNAs from a
CC   biological sample, in expression profiling studies, in qualitatively or
CC   quantitatively characterising the corresponding transcriptome, and in
CC   detecting RNA transcripts and splice variants of human or animal
CC   transcriptomes. The libraries may also be used as specialised mini
CC   libraries to detect transcripts of a sub-transcriptome under a
CC   particular biological or pathological state, and so allowing the
CC   detection of tissue- and pathology-specific genes such as those genes
CC   only expressed in specific tissue under a specific pathological
CC   condition; to detect developmental specific genes; and to detect RNA
CC   transcripts and splice variants of a transcriptome of a patient suffering
CC   from a particular disorder. ABN27253 to ABN59589 represent
CC   oligonucleotide sequences from rats, humans and mice, which are used in
CC   the exemplification of the present invention.
CC   N.B. The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 60 BP; 19 A; 10 C; 16 G; 15 T; 0 other;

  Query Match             72.0%;  Score 14.4;  DB 24;  Length 60;
  Best Local Similarity   93.8%;  Pred. No. 1.4e+03;
  Matches   15;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 GCTGAGAAACCTTCTT 16
           ||| ||||||||||||
Db      35 GCTCAGAAACCTTCTT 20
```

```
RESULT 11
AAT66259
ID   AAT66259 standard; DNA; 20 BP.
XX
AC   AAT66259;
XX
DT   27-DEC-1997  (first entry)
XX
DE   Primer 3 for hop gene.
XX
KW   primer; PCR; polymerase chain reaction; amplification; hop;
KW   polymorphism; determination; analysis; genetic variation; ss.
XX
OS   Synthetic.
XX
PN   WO9705281-A1.
XX
PD   13-FEB-1997.
XX
PF   26-JUL-1996;   96WO-JP02121.
XX
PR   30-APR-1996;   96JP-0130586.
PR   28-JUL-1995;   95JP-0211328.
XX
PA   (SAPB ) SAPPORO BREWERIES.
XX
PI   Araki S,  Tsuchiya Y;
XX
DR   WPI; 1997-145715/13.
XX
PT   Amplifying the polymorphic region in hop DNA using specific primers
PT   - useful for distinguishing between varieties of hops
XX
PS   Claim 6; Page 32; 58pp; Japanese.
XX
CC   AAT66257-96 are primers used for PCR amplification of a hop gene
CC   containing an intervarietal polymorphism. Different varieties of hops
CC   can be distinguished genetically by conducting PCR on a DNA sample from
CC   the hops using these primers, and then analysing the amplification
CC   products (e.g. by restriction enzyme cleavage). The use of several
CC   different primers allows the genetic variation to be studied in detail.
XX
SQ   Sequence 20 BP; 4 A; 4 C; 6 G; 6 T; 0 other;

  Query Match             71.0%;  Score 14.2;  DB 18;  Length 20;
  Best Local Similarity   84.2%;  Pred. No. 1.6e+03;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 GCTGAGAAACCTTCTTTTG 19
           |||||| || ||||||| |
Db       2 GCTGAGCAAGCTTCTTTGG 20
```

```
RESULT 12
AAX88476
ID   AAX88476 standard; DNA; 26 BP.
XX
AC   AAX88476;
XX
DT   01-OCT-1999  (first entry)
XX
DE   Human MIP-1 beta primer P2C.
XX
KW   RANTES; chemokine; detection; primer; probe; amplification; MIP-1 alpha;
KW   regulated upon activation normal T expressed and secreted; MIP-1 beta;
KW   macrophage inflammatory protein; CD4+ T-cell; inhibitor; prognosis;
KW   primary non-syncytium-inducing HIV-1 strain; therapy; ss.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
```

PN WO9937815-A1.
XX
PD 29-JUL-1999.
XX
PF 22-JAN-1999; 99WO-US01327.
XX
PR 22-JAN-1998; 98US-0010641.
XX
PA (ALKU) AKZO NOBEL NV.
XX
PI Romano JW, Shurtliff R, Williams KG;
XX
DR WPI; 1999-469145/39.
XX
PT Detection of expression levels of the cytokines RANTES, MIP-1alpha
PT and MIP-1beta used as prognostic markers of HIV-infected patients
XX
PS Claim 1; Page 42; 48pp; English.
XX
CC This invention describes novel oligonucleotides which are used for
CC detecting the chemokines RANTES (regulated upon activation normal T
CC expressed and secreted), macrophage inflammatory protein (MIP)-1 alpha
CC or MIP-1 beta by (a) obtaining a sample possible containing RANTES or
CC MIP-1 alpha or MIP-1 beta RNA, (b) performing an isothermal
CC transcriptional amplification on the sample with 2 oligonucleotide
CC primers, (c) detecting the product of step (b) where detection of a
CC product indicates the presence of RANTES, MIP-1 alpha or MIP-1 beta in
CC the sample. The assay is used to determine the levels of the chemokines
CC RANTES, MIP-1 alpha and MIP-1 beta in samples, especially cells. These
CC chemokines have been shown to be inhibitors of CD4+ T-cells by primary
CC non-syncytium-inducing HIV-1 strains. Thus the level of expression of
CC these genes can be used as prognostic markers for direct therapeutic
CC management of HIV-infected patients. By being isothermic, the assay
CC requires less manipulation by the experimenter. Also 'spiking' the
CC sample with a known amount of control RNA allows quantitation and
CC qualification of the products in a single assay. AAX88447-X88491
CC represent the primers and probes used inthe method of the invention.
XX
SQ Sequence 26 BP; 7 A; 9 C; 2 G; 8 T; 0 other;

Query Match 71.0%; Score 14.2; DB 20; Length 26;
Best Local Similarity 84.2%; Pred. No. 1.6e+03;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 CTGAGAAACCTTCTTTTGC 20
||||||||| | |||||| |
Db 7 CTGAGAAAACCTCTTTTCC 25

RESULT 13
AAD00984
ID AAD00984 standard; DNA; 29 BP.
XX
AC AAD00984;
XX
DT 21-SEP-2000 (first entry)
XX
DE Primer #2 for AtEM1 CAPS marker to map Arabidopsis thaliana PAD4 gene.
XX
KW PAD4; disease resistance; phytoalexin; PR-1; PR-5;
KW pathogenesis-related protein; BGL2; beta-glucanase; ASA1;
KW anthranilate synthase; defence response; salicylic acid; SA;
KW signal transduction; transgenic plant; pathogen; bacteria; fungi;
KW nematode; Phytophthora; Peronospora; Pseudomonas; plant; agronomy;
KW crop; chromosome 3; pad4-1; PCR primer; AtEM1 CAPS marker;
KW Cleaved Amplified Polymorphic Sequence marker; ss.
XX
OS Arabidopsis thaliana.
XX
PN WO200029595-A1.
XX
PD 25-MAY-2000.
XX
PF 04-NOV-1999; 99WO-US26106.
XX
PR 12-NOV-1998; 98US-0190733.
XX
PA (UYMA-) UNIV MARYLAND BIOTECHNOLOGY INST.
PA (PLAN-) PLANT BIOSCIENCE LTD.
XX
PI Glazebrook J, Jirage D, Tootle T, Feys BJF;
XX
DR WPI; 2000-387805/33.
XX
PT New PAD4 polypeptide from Arabidopsis thaliana, useful to enhance plant
PT resistance to diseases due to pathogens such as Phytophthora e.g. to
PT improve crop quality or yields -
XX
PS Example 2; Page 48; 181pp; English.
XX
CC The present sequence is a primer for AtEM1 Cleaved Amplified
CC Polymorphic Sequence (CAPS) marker which is used to map
CC Arabidopsis thaliana PAD4 and pad4-1 genes.
CC PAD4 gene is located on Arabidopsis chromosome 3 and encodes a
CC protein which plays an important role in disease resistance
CC in plants. The protein has positive
CC regulatory effect on phytoalexin levels and PR-1 (pathogenesis-related
CC protein) expression levels, but has no effect on PR-5 (pathogenesis-
CC related protein), BGL2 (beta-glucanase) or ASA1 (anthranilate synthase)
CC expression levels in a disease defence response by a host plant.
CC PAD4 is required upstream from salicylic acid in the signal
CC transduction pathway leading from infection to activation of defence
CC responses. It is used to produce transgenic plants which have enhanced
CC resistance to diseases caused due to pathogens such as bacteria, fungi,
CC and nematodes, especially Phytophthora, Peronospora or Pseudomonas.
CC Such transgenic plants are useful agronomically e.g. to improve crop
CC quality or yield.
XX
SQ Sequence 29 BP; 6 A; 6 C; 8 G; 9 T; 0 other;

Query Match 71.0%; Score 14.2; DB 21; Length 29;
Best Local Similarity 84.2%; Pred. No. 1.6e+03;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 CTGAGAAACCTTCTTTTGC 20
||| ||||||||| || |||
Db 4 CTGGGAAACCTTATTATGC 22

RESULT 14
AAQ48439
ID AAQ48439 standard; DNA; 20 BP.
XX
AC AAQ48439;
XX
DT 23-MAR-1994 (first entry)
XX
DE Brh-I N-terminal primer #1.
XX
KW Wasp; insect; toxicity; ectoparasite; Habrobracon hebetor; venom; PCR;
KW Microbracon hebetor; paralysis; Lepidoptera; polymerase chain reaction;
KW primer; amplify; Bracon hebetor; ss.
XX
OS Synthetic.
XX
PN WO9318145-A.
XX
PD 16-SEP-1993.
XX
PF 25-FEB-1993; 93WO-EP00431.
XX
PR 04-MAR-1992; 92US-0847570.
PR 10-JUN-1992; 92US-0897192.
XX
PA (SANO) SANDOZ LTD.
PA (SANO) SANDOZ PATENT GMBH.

```
PA   (SANO ) SANDOZ-ERFINDUNGEN VERW GES MBH.
XX
PI   Leisy DJ,  Quistad GB;
XX
DR   WPI; 1993-303459/38.
XX
PT   Polypeptide(s) from Bracon hebetor venom - useful as insecticides,
PT   when cloned into baculoviruses
XX
PS   Example 5; Page 12; 43pp; English.
XX
CC   The sequences given in AAQ48439-40 are primers which were used in the
CC   synthesis of single-stranded cDNA encoding the Bracon hebetor toxin,
CC   Brh-I.  Primer #1 is complementary to a sequence derived by reverse
CC   translation of the Brh-I toxin N-terminal.  Primer #2 corresponds to a
CC   second portion of the reverse translation product of the Brh-I N-
CC   terminal.  The amplification product encodes a wasp polypeptide which
CC   exhibits insect toxicity and has a molecular weight >70 kD.  This
CC   sequence was derived from the ectoparasitic wasp, also known as
CC   Habrobracon hebetor and Microbracon hebetor.  The protein encoded
CC   by the amplification product is a venom which is paralytic to
CC   Lepidopterans, and may therefore be used for controlling insects,
CC   esp. Lepidoptera.  This protein is quite labile under cetain isolation
CC   conditions, it is particularly unstable at low pH.
XX
SQ   Sequence 20 BP; 4 A; 1 C; 2 G; 6 T; 7 other;

  Query Match             70.0%;  Score 14;  DB 14;  Length 20;
  Best Local Similarity   57.9%;  Pred. No. 2e+03;
  Matches   11;  Conservative    6;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 GCTGAGAAACCTTCTTTTG 19
           ::||: | ||:||:||:|| 
Db       2 RYTGRAANACYTTYTTYTG 20


RESULT 15
AAT45676
ID   AAT45676 standard; DNA; 20 BP.
XX
AC   AAT45676;
XX
DT   07-FEB-1997  (first entry)
XX
DE   Bracon hebetor Brh-1 toxin gene, degenerate RT-PCR primer.
XX
KW   Neurotoxin; toxic; pesticide; insecticide; paralysis; venom gland;
KW   lepidoptera; spodoptera; reverse transcriptase; RT;
KW   polymerase chain reaction; PCR; ss.
XX
OS   Bracon hebetor.
XX
PN   US5554592-A.
XX
PD   10-SEP-1996.
XX
PF   04-MAR-1992;   92US-0847570.
XX
PR   05-FEB-1993;   93US-0013890.
PR   04-MAR-1992;   92US-0847570.
PR   10-JUN-1992;   92US-0897192.
PR   10-AUG-1994;   94US-0288408.
XX
PA   (SANO ) SANDOZ LTD.
XX
PI   Leisy DJ,  Quistad GB;
XX
DR   WPI; 1996-424679/42.
XX
PT   New isolated Bracon hebetor wasp polypeptide(s) - useful as
PT   insecticides which are toxic to insects, partic. of the order
PT   lepidoptera
XX
PS   Example 5; Column 8; 23pp; English.
XX
CC   AAT45676 and AAT45677 are degenerate primers used in a RT-PCR for the
CC   isolation of the Brh-1 toxin gene sequence derived from mRNA of
CC   the venom glands of the parasitic wasp, Bracon hebetor (Brh).
CC   The toxin causes paralysis and death in insects of the order
CC   Lepidoptera, e.g. Heliothis viscerens (tobacco budworm, evidence
CC   given in the specification) and of the order Spodoptera. The toxin
CC   is useful as an insecticide to protect plants (esp. crops) from
CC   damage caused by insect pests. The Brh-1 toxin has a mol. wt. of
CC   greater than 70 kD.
XX
SQ   Sequence 20 BP; 4 A; 1 C; 2 G; 6 T; 7 other;

  Query Match             70.0%;  Score 14;  DB 17;  Length 20;
  Best Local Similarity   57.9%;  Pred. No. 2e+03;
  Matches   11;  Conservative    6;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 GCTGAGAAACCTTCTTTTG 19
           ::||: | ||:||:||:|| 
Db       2 RYTGRAANACYTTYTTYTG 20


Search completed: November 23, 2002, 07:03:38
Job time : 98.1 secs
```

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 06:36:31 ; Search time 21.3 Seconds
(without alignments)
287.959 Million cell updates/sec

Title: US-09-296-264-21
Perfect score: 20
Sequence: 1 gctgagaaaccttcttttgc 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 687286

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	16.4	82.0	68	1	US-08-288-408-11	Sequence 11, Appl
c	2	16.4	82.0	68	4	US-08-655-782-11	Sequence 11, Appl
c	3	15.2	76.0	68	1	US-08-288-408-13	Sequence 13, Appl
c	4	15.2	76.0	68	4	US-08-655-782-13	Sequence 13, Appl
c	5	14.6	73.0	93	1	US-08-288-408-10	Sequence 10, Appl
c	6	14.6	73.0	93	4	US-08-655-782-10	Sequence 10, Appl
	7	14.2	71.0	20	2	US-08-809-297-3	Sequence 3, Appli
	8	14.2	71.0	26	3	US-09-010-641-30	Sequence 30, Appl
	9	14.2	71.0	26	4	US-09-356-281-30	Sequence 30, Appl
	10	14	70.0	20	1	US-08-288-408-6	Sequence 6, Appli
	11	14	70.0	20	4	US-08-655-782-6	Sequence 6, Appli
	12	13.6	68.0	32	1	US-08-417-330A-7	Sequence 7, Appli
c	13	13.6	68.0	68	1	US-08-288-408-12	Sequence 12, Appl
c	14	13.6	68.0	68	4	US-08-655-782-12	Sequence 12, Appl
c	15	12.8	64.0	29	2	US-08-480-736-2	Sequence 2, Appli
c	16	12.8	64.0	49	1	US-08-171-389-118	Sequence 118, App
c	17	12.8	64.0	49	1	US-08-123-936-118	Sequence 118, App
c	18	12.8	64.0	49	2	US-08-475-228A-118	Sequence 118, App
c	19	12.8	64.0	49	3	US-08-482-080A-118	Sequence 118, App
c	20	12.8	64.0	49	4	US-09-354-947-118	Sequence 118, App
c	21	12.8	64.0	49	5	PCT-US93-12388-118	Sequence 118, App
c	22	12.6	63.0	20	4	US-09-021-701-239	Sequence 239, App
c	23	12.6	63.0	20	4	US-09-021-701-240	Sequence 240, App
c	24	12.6	63.0	43	3	US-08-732-708C-12	Sequence 12, Appl
	25	12.6	63.0	47	4	US-09-641-638-801	Sequence 801, App
	26	12.6	63.0	57	4	US-09-037-990B-38	Sequence 38, Appl
	27	12.6	63.0	59	4	US-09-037-990B-40	Sequence 40, Appl
	28	12.6	63.0	60	1	US-07-960-932-2	Sequence 2, Appli
c	29	12.6	63.0	60	1	US-07-960-932-3	Sequence 3, Appli
	30	12.6	63.0	60	1	US-08-455-970A-3	Sequence 3, Appli
c	31	12.6	63.0	60	1	US-08-455-970A-4	Sequence 4, Appli
	32	12.6	63.0	61	4	US-09-290-577-60	Sequence 60, Appl
	33	12.6	63.0	61	4	US-09-290-452-60	Sequence 60, Appl
	34	12.6	63.0	61	4	US-09-290-338-60	Sequence 60, Appl
	35	12.4	62.0	17	4	US-08-584-040-1984	Sequence 1984, Ap
	36	12.4	62.0	20	2	US-08-500-857A-12	Sequence 12, Appl
c	37	12.4	62.0	36	4	US-08-250-802-24	Sequence 24, Appl
c	38	12.4	62.0	36	5	PCT-US92-07916-24	Sequence 24, Appl
	39	12.4	62.0	38	4	US-08-250-802-27	Sequence 27, Appl
	40	12.4	62.0	38	5	PCT-US92-07916-27	Sequence 27, Appl
	41	12.4	62.0	77	1	US-08-399-412A-35	Sequence 35, Appl
c	42	12.2	61.0	19	1	US-08-832-883-16	Sequence 16, Appl
c	43	12.2	61.0	19	2	US-08-832-877-16	Sequence 16, Appl
	44	12.2	61.0	30	6	5202236-27	Patent No. 5202236
c	45	12.2	61.0	33	1	US-08-195-874-7	Sequence 7, Appli

ALIGNMENTS

```
RESULT 1
US-08-288-408-11/c
; Sequence 11, Application US/08288408
; Patent No. 5554592
;  GENERAL INFORMATION:
;    APPLICANT:  Quistad, Gary B.
;    APPLICANT:  Leisy, Douglas J.
;    TITLE OF INVENTION:  INSECTICIDAL TOXINS FROM THE PARASITIC
;    TITLE OF INVENTION:  WASP BRACON HEBETOR
;    NUMBER OF SEQUENCES:  13
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sandoz Agro, Inc.
;      STREET:  975 California Avenue
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304-1104
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/288,408
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/013,890
;      FILING DATE:  05-FEB-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/897,192
;      FILING DATE:  10-JUN-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/847,570
;      FILING DATE:  04-MAR-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  No. 5554592ris, Allen E.
;      REGISTRATION NUMBER:  34,490
;      REFERENCE/DOCKET NUMBER:
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415/354-3592
;      TELEFAX:  415/857-1125
;  INFORMATION FOR SEQ ID NO:  11:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  68 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-288-408-11
```

```
  Query Match             82.0%;  Score 16.4;  DB 1;  Length 68;
  Best Local Similarity   94.4%;  Pred. No. 11;
  Matches   17;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      3 TGAGAAACCTTCTTTTGC 20
          ||| ||||||||||||||
Db     65 TGAAAAACCTTCTTTTGC 48


RESULT 2
US-08-655-782-11/c
; Sequence 11, Application US/08655782
; Patent No. 6251862
;  GENERAL INFORMATION:
;    APPLICANT:  Quistad, Gary B.
;    APPLICANT:  Leisy, Douglas J.
;    TITLE OF INVENTION:  INSECTICIDAL TOXINS FROM THE PARASITIC
;    TITLE OF INVENTION:  WASP BRACON HEBETOR
;    NUMBER OF SEQUENCES:  13
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sandoz Agro, Inc.
;      STREET:  975 California Avenue
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304-1104
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/655,782
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;    PRIOR APPLICATION DATA:  A) APPLICATION NUMBER: US/08/288,408
;      FILING DATE:
;      APPLICATION NUMBER:  US 08/013,890
;      FILING DATE:  05-FEB-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/897,192
;      FILING DATE:  10-JUN-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/847,570
;      FILING DATE:  04-MAR-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  No. 6251862ris, Allen E.
;      REGISTRATION NUMBER:  34,490
;      REFERENCE/DOCKET NUMBER:
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415/354-3592
;      TELEFAX:  415/857-1125
;  INFORMATION FOR SEQ ID NO:  11:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  68 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-655-782-11

  Query Match             82.0%;  Score 16.4;  DB 4;  Length 68;
  Best Local Similarity   94.4%;  Pred. No. 11;
  Matches   17;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      3 TGAGAAACCTTCTTTTGC 20
          ||| ||||||||||||||
Db     65 TGAAAAACCTTCTTTTGC 48


RESULT 3
US-08-288-408-13/c
; Sequence 13, Application US/08288408
; Patent No. 5554592
;  GENERAL INFORMATION:
;    APPLICANT:  Quistad, Gary B.
;    APPLICANT:  Leisy, Douglas J.
;    TITLE OF INVENTION:  INSECTICIDAL TOXINS FROM THE PARASITIC
;    TITLE OF INVENTION:  WASP BRACON HEBETOR
;    NUMBER OF SEQUENCES:  13
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sandoz Agro, Inc.
;      STREET:  975 California Avenue
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304-1104
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/288,408
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/013,890
;      FILING DATE:  05-FEB-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/897,192
;      FILING DATE:  10-JUN-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/847,570
;      FILING DATE:  04-MAR-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  No. 5554592ris, Allen E.
;      REGISTRATION NUMBER:  34,490
;      REFERENCE/DOCKET NUMBER:
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415/354-3592
;      TELEFAX:  415/857-1125
;  INFORMATION FOR SEQ ID NO:  13:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  68 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-288-408-13

  Query Match             76.0%;  Score 15.2;  DB 1;  Length 68;
  Best Local Similarity   85.0%;  Pred. No. 45;
  Matches   17;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 GCTGAGAAACCTTCTTTTGC 20
          | ||| | ||||||||||||
Db     67 GTTGAAATACCTTCTTTTGC 48


RESULT 4
US-08-655-782-13/c
; Sequence 13, Application US/08655782
; Patent No. 6251862
;  GENERAL INFORMATION:
;    APPLICANT:  Quistad, Gary B.
;    APPLICANT:  Leisy, Douglas J.
;    TITLE OF INVENTION:  INSECTICIDAL TOXINS FROM THE PARASITIC
;    TITLE OF INVENTION:  WASP BRACON HEBETOR
;    NUMBER OF SEQUENCES:  13
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sandoz Agro, Inc.
;      STREET:  975 California Avenue
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304-1104
```

```
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/655,782
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;    PRIOR APPLICATION DATA:  A) APPLICATION NUMBER: US/08/288,408
;      FILING DATE:
;      APPLICATION NUMBER:  US 08/013,890
;      FILING DATE:  05-FEB-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/897,192
;      FILING DATE:  10-JUN-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/847,570
;      FILING DATE:  04-MAR-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  No. 6251862ris, Allen E.
;      REGISTRATION NUMBER:  34,490
;      REFERENCE/DOCKET NUMBER:
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415/354-3592
;      TELEFAX:  415/857-1125
;  INFORMATION FOR SEQ ID NO:  13:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  68 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-655-782-13

  Query Match             76.0%;  Score 15.2;  DB 4;  Length 68;
  Best Local Similarity   85.0%;  Pred. No. 45;
  Matches   17;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 GCTGAGAAACCTTCTTTTGC 20
           | ||| | ||||||||||||
Db      67 GTTGAAATACCTTCTTTTGC 48


RESULT 5
US-08-288-408-10/c
; Sequence 10, Application US/08288408
; Patent No. 5554592
;  GENERAL INFORMATION:
;    APPLICANT:  Quistad, Gary B.
;    APPLICANT:  Leisy, Douglas J.
;    TITLE OF INVENTION:  INSECTICIDAL TOXINS FROM THE PARASITIC
;    TITLE OF INVENTION:  WASP BRACON HEBETOR
;    NUMBER OF SEQUENCES:  13
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sandoz Agro, Inc.
;      STREET:  975 California Avenue
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304-1104
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/288,408
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/013,890
;      FILING DATE:  05-FEB-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/897,192
;      FILING DATE:  10-JUN-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/847,570
;      FILING DATE:  04-MAR-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  No. 5554592ris, Allen E.
;      REGISTRATION NUMBER:  34,490
;      REFERENCE/DOCKET NUMBER:
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415/354-3592
;      TELEFAX:  415/857-1125
;  INFORMATION FOR SEQ ID NO:  10:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  93 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-288-408-10

  Query Match             73.0%;  Score 14.6;  DB 1;  Length 93;
  Best Local Similarity   55.0%;  Pred. No. 96;
  Matches   11;  Conservative    7;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 GCTGAGAAACCTTCTTTTGC 20
           ::||: | ||:||:||:||:
Db      70 RYTGRAANACYTTYTTYTGY 51


RESULT 6
US-08-655-782-10/c
; Sequence 10, Application US/08655782
; Patent No. 6251862
;  GENERAL INFORMATION:
;    APPLICANT:  Quistad, Gary B.
;    APPLICANT:  Leisy, Douglas J.
;    TITLE OF INVENTION:  INSECTICIDAL TOXINS FROM THE PARASITIC
;    TITLE OF INVENTION:  WASP BRACON HEBETOR
;    NUMBER OF SEQUENCES:  13
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sandoz Agro, Inc.
;      STREET:  975 California Avenue
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304-1104
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/655,782
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;    PRIOR APPLICATION DATA:  A) APPLICATION NUMBER: US/08/288,408
;      FILING DATE:
;      APPLICATION NUMBER:  US 08/013,890
;      FILING DATE:  05-FEB-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/897,192
;      FILING DATE:  10-JUN-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/847,570
;      FILING DATE:  04-MAR-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  No. 6251862ris, Allen E.
;      REGISTRATION NUMBER:  34,490
;      REFERENCE/DOCKET NUMBER:
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415/354-3592
```

```
;      TELEFAX:  415/857-1125
;  INFORMATION FOR SEQ ID NO:  10:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  93 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-655-782-10

  Query Match             73.0%;  Score 14.6;  DB 4;  Length 93;
  Best Local Similarity   55.0%;  Pred. No. 96;
  Matches   11;  Conservative    7;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 GCTGAGAAACCTTCTTTTGC 20
          ::||: | ||:||:||:||:
Db     70 RYTGRAANACYTTYTTYTGY 51


RESULT 7
US-08-809-297-3
; Sequence 3, Application US/08809297
; Patent No. 5948650
;  GENERAL INFORMATION:
;    APPLICANT:  ARAKI, SHIGEKI
;    APPLICANT:  TSUCHIYA, YOHICHI
;    TITLE OF INVENTION:  GENETIC VARIETY IDENTIFYING METHOD IN
;    TITLE OF INVENTION:  HOPS
;    NUMBER OF SEQUENCES:  48
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
;      ADDRESSEE:  P.C.
;      STREET:  1755 SOUTH JEFFERSON DAVIS HIGHWAY, SUITE 400
;      CITY:  ARLINGTON
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22202
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/809,297
;      FILING DATE:  06-MAY-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/JP96/02121
;      FILING DATE:  26-JUL-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP HEI 7-211328
;      FILING DATE:  28-JUL-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP HEI 8-130586
;      FILING DATE:  30-APR-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  OBLON, NORMAN F.
;      REGISTRATION NUMBER:  24618
;      REFERENCE/DOCKET NUMBER:  2589-057-0PCT
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703) 413-3000
;      TELEFAX:  (703) 413-2220
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  other nucleic acid
;      DESCRIPTION:  /desc = "SYNTHETIC DNA"
US-08-809-297-3

  Query Match             71.0%;  Score 14.2;  DB 2;  Length 20;
  Best Local Similarity   84.2%;  Pred. No. 1.2e+02;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 GCTGAGAAACCTTCTTTTG 19
          |||||| || ||||||| |
Db      2 GCTGAGCAAGCTTCTTTGG 20


RESULT 8
US-09-010-641-30
; Sequence 30, Application US/09010641
; Patent No. 6121023
;  GENERAL INFORMATION:
;    APPLICANT:  ROMANO, JOSEPH W.
;    APPLICANT:  SHURTLIFF, ROXANNE
;    APPLICANT:  WILLIAMS, KIMBERLY G.
;    TITLE OF INVENTION:  ISOTHERMAL AMPLIFICATION BASED ASSAY FOR
;    TITLE OF INVENTION:  THE DETECTION AND QUANTIFICATION OF CHEMOKINES RANTES,
;    TITLE OF INVENTION:  MIP-1ALPHA AND MIP-1BETA
;    NUMBER OF SEQUENCES:  45
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  AKZO NOBEL PATENT DEPARTMENT
;      STREET:  1300 PICCARD DRIVE, SUITE 206
;      CITY:  ROCKVILLE
;      STATE:  MARYLAND
;      COUNTRY:  USA
;      ZIP:  20850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/010,641
;      FILING DATE:  22-JAN-1998
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  KLESNER, SHARON N.
;      REGISTRATION NUMBER:  36,335
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  301-948-7400
;      TELEFAX:  301-948-9751
;  INFORMATION FOR SEQ ID NO:  30:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  26 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA
US-09-010-641-30

  Query Match             71.0%;  Score 14.2;  DB 3;  Length 26;
  Best Local Similarity   84.2%;  Pred. No. 1.2e+02;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 CTGAGAAACCTTCTTTTGC 20
          |||||||| | |||||| |
Db      7 CTGAGAAAACCTCTTTTCC 25


RESULT 9
US-09-356-281-30
; Sequence 30, Application US/09356281
; Patent No. 6218154
;  GENERAL INFORMATION:
;    APPLICANT:  ROMANO, JOSEPH W.
;    APPLICANT:  SHURTLIFF, ROXANNE
;    APPLICANT:  WILLIAMS, KIMBERLY G.
;    TITLE OF INVENTION:  ISOTHERMAL AMPLIFICATION BASED ASSAY FOR
;    TITLE OF INVENTION:  THE DETECTION AND QUANTIFICATION OF CHEMOKINES RANTES,
;    TITLE OF INVENTION:  MIP-1ALPHA AND MIP-1BETA
;    NUMBER OF SEQUENCES:  45
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  AKZO NOBEL PATENT DEPARTMENT
```

```
;       STREET:  1300 PICCARD DRIVE, SUITE 206
;       CITY:  ROCKVILLE
;       STATE:  MARYLAND
;       COUNTRY:  USA
;       ZIP:  20850
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/09/356,281
;       FILING DATE:
;       CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  09/010,641
;       FILING DATE:  22-JAN-1998
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  KLESNER, SHARON N.
;       REGISTRATION NUMBER:  36,335
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  301-948-7400
;       TELEFAX:  301-948-9751
;   INFORMATION FOR SEQ ID NO:  30:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  26 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA
US-09-356-281-30

  Query Match             71.0%;  Score 14.2;  DB 4;  Length 26;
  Best Local Similarity   84.2%;  Pred. No. 1.2e+02;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 CTGAGAAACCTTCTTTTGC 20
           |||||||| | |||||| |
Db       7 CTGAGAAAACCTCTTTTCC 25


RESULT 10
US-08-288-408-6
; Sequence 6, Application US/08288408
; Patent No. 5554592
;  GENERAL INFORMATION:
;    APPLICANT:  Quistad, Gary B.
;    APPLICANT:  Leisy, Douglas J.
;    TITLE OF INVENTION:  INSECTICIDAL TOXINS FROM THE PARASITIC
;    TITLE OF INVENTION:  WASP BRACON HEBETOR
;    NUMBER OF SEQUENCES:  13
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sandoz Agro, Inc.
;      STREET:  975 California Avenue
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304-1104
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/288,408
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/013,890
;      FILING DATE:  05-FEB-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/897,192
;      FILING DATE:  10-JUN-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/847,570
;      FILING DATE:  04-MAR-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  No. 5554592ris, Allen E.
;      REGISTRATION NUMBER:  34,490
;      REFERENCE/DOCKET NUMBER:
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415/354-3592
;      TELEFAX:  415/857-1125
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-288-408-6

  Query Match             70.0%;  Score 14;  DB 1;  Length 20;
  Best Local Similarity   57.9%;  Pred. No. 1.5e+02;
  Matches   11;  Conservative    6;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 GCTGAGAAACCTTCTTTTG 19
           ::||: | ||:||:||:||
Db       2 RYTGRAANACYTTYTTYTG 20


RESULT 11
US-08-655-782-6
; Sequence 6, Application US/08655782
; Patent No. 6251862
;  GENERAL INFORMATION:
;    APPLICANT:  Quistad, Gary B.
;    APPLICANT:  Leisy, Douglas J.
;    TITLE OF INVENTION:  INSECTICIDAL TOXINS FROM THE PARASITIC
;    TITLE OF INVENTION:  WASP BRACON HEBETOR
;    NUMBER OF SEQUENCES:  13
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sandoz Agro, Inc.
;      STREET:  975 California Avenue
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304-1104
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/655,782
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;    PRIOR APPLICATION DATA:  A) APPLICATION NUMBER: US/08/288,408
;      FILING DATE:
;      APPLICATION NUMBER:  US 08/013,890
;      FILING DATE:  05-FEB-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/897,192
;      FILING DATE:  10-JUN-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/847,570
;      FILING DATE:  04-MAR-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  No. 6251862ris, Allen E.
;      REGISTRATION NUMBER:  34,490
;      REFERENCE/DOCKET NUMBER:
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415/354-3592
;      TELEFAX:  415/857-1125
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
```

```
;      LENGTH:  20 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-655-782-6

  Query Match           70.0%;  Score 14;  DB 4;  Length 20;
  Best Local Similarity   57.9%;  Pred. No. 1.5e+02;
  Matches   11;  Conservative    6;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 GCTGAGAAACCTTCTTTTG 19
          ::||: | ||:||:||:||
Db      2 RYTGRAANACYTTYTTYTG 20


RESULT 12
US-08-417-330A-7
; Sequence 7, Application US/08417330A
; Patent No. 5719057
;  GENERAL INFORMATION:
;    APPLICANT:  HADINGHAM, KAREN
;    APPLICANT:  LE BOURDELLES, BEATRICE
;    APPLICANT:  WHITING, PAUL
;    APPLICANT:  WINGROVE, PETER
;    TITLE OF INVENTION:  STABLY TRANSFECTED CELL LINE EXPRESSING
;    TITLE OF INVENTION:  GABA-A RECEPTOR AND NOVEL CLONED
;    TITLE OF INVENTION:  GABA-RECEPTOR SUBUNIT cDNA SEQUENCES
;    NUMBER OF SEQUENCES:  20
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  J. MARK HAND - MERCK & CO., INC.
;      STREET:  126 EAST LINCOLN AVENUE - P.O. BOX 2000
;      CITY:  RAHWAY
;      STATE:  NJ
;      COUNTRY:  US
;      ZIP:  07065-0900
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/417,330A
;      FILING DATE:  05-APR-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  HAND, MARK
;      REGISTRATION NUMBER:  36,545
;      REFERENCE/DOCKET NUMBER:  T1109IA
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  908-594-3905
;      TELEFAX:  908-594-4720
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  32 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-417-330A-7

  Query Match           68.0%;  Score 13.6;  DB 1;  Length 32;
  Best Local Similarity   80.0%;  Pred. No. 2.6e+02;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 GCTGAGAAACCTTCTTTTGC 20
          |||||||||  ||  ||  |||
Db     13 GCTGAGAAAGCTGCTAATGC 32


RESULT 13
US-08-288-408-12/c
; Sequence 12, Application US/08288408
; Patent No. 5554592
;  GENERAL INFORMATION:
;    APPLICANT:  Quistad, Gary B.
;    APPLICANT:  Leisy, Douglas J.
;    TITLE OF INVENTION:  INSECTICIDAL TOXINS FROM THE PARASITIC
;    TITLE OF INVENTION:  WASP BRACON HEBETOR
;    NUMBER OF SEQUENCES:  13
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sandoz Agro, Inc.
;      STREET:  975 California Avenue
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304-1104
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/288,408
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/013,890
;      FILING DATE:  05-FEB-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/897,192
;      FILING DATE:  10-JUN-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/847,570
;      FILING DATE:  04-MAR-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  No. 5554592ris, Allen E.
;      REGISTRATION NUMBER:  34,490
;      REFERENCE/DOCKET NUMBER:
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415/354-3592
;      TELEFAX:  415/857-1125
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  68 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-288-408-12

  Query Match           68.0%;  Score 13.6;  DB 1;  Length 68;
  Best Local Similarity   80.0%;  Pred. No. 2.9e+02;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 GCTGAGAAACCTTCTTTTGC 20
          | ||| | ||||| ||||||
Db     67 GTTGAAATACCTTTTTTTGC 48


RESULT 14
US-08-655-782-12/c
; Sequence 12, Application US/08655782
; Patent No. 6251862
;  GENERAL INFORMATION:
;    APPLICANT:  Quistad, Gary B.
;    APPLICANT:  Leisy, Douglas J.
;    TITLE OF INVENTION:  INSECTICIDAL TOXINS FROM THE PARASITIC
;    TITLE OF INVENTION:  WASP BRACON HEBETOR
;    NUMBER OF SEQUENCES:  13
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sandoz Agro, Inc.
;      STREET:  975 California Avenue
;      CITY:  Palo Alto
```

```
;       STATE:  CA
;       COUNTRY:  USA
;       ZIP:  94304-1104
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/655,782
;       FILING DATE:
;       CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;     PRIOR APPLICATION DATA:  A) APPLICATION NUMBER: US/08/288,408
;       FILING DATE:
;       APPLICATION NUMBER:  US 08/013,890
;       FILING DATE:  05-FEB-1993
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/897,192
;       FILING DATE:  10-JUN-1992
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/847,570
;       FILING DATE:  04-MAR-1992
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  No. 6251862ris, Allen E.
;       REGISTRATION NUMBER:  34,490
;       REFERENCE/DOCKET NUMBER:
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  415/354-3592
;       TELEFAX:  415/857-1125
;   INFORMATION FOR SEQ ID NO:  12:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  68 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
US-08-655-782-12

  Query Match             68.0%;  Score 13.6;  DB 4;  Length 68;
  Best Local Similarity   80.0%;  Pred. No. 2.9e+02;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 GCTGAGAAACCTTCTTTTGC 20
           | ||| | ||||| ||||||
Db      67 GTTGAAATACCTTTTTTTGC 48


RESULT 15
US-08-480-736-2/c
; Sequence 2, Application US/08480736
; Patent No. 5830477
;  GENERAL INFORMATION:
;     APPLICANT:  LATHE, Richard
;     APPLICANT:  KIENY, Marie-Paule
;     APPLICANT:  DRILLIEN, Robert
;     APPLICANT:  LECOCQ, Jean-Pierre
;     TITLE OF INVENTION:  VACCINE AGAINST RABIES AND PROCESS FOR
;     TITLE OF INVENTION:  PREPARATION THEREOF
;     NUMBER OF SEQUENCES:  12
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Burns, Doane, Swecker & Mathis
;       STREET:  P.O. Box 1404
;       CITY:  Alexandria
;       STATE:  Virginia
;       COUNTRY:  United States
;       ZIP:  22313-1404
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/480,736
;       FILING DATE:  07-JUN-1995
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/231,457
;       FILING DATE:  21-APR-1994
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/038,052
;       FILING DATE:  29-MAR-1993
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/759,138
;       FILING DATE:  11-SEP-1991
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 07/378,801
;       FILING DATE:  11-JUL-1989
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 06/829,144
;       FILING DATE:  24-DEC-1985
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  FR 84/06499
;       FILING DATE:  25-APR-1984
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  WO PCT/FR85/00096
;       FILING DATE:  24-APR-1985
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Rea, Teresa Stanek
;       REGISTRATION NUMBER:  30,427
;       REFERENCE/DOCKET NUMBER:  017753-061
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (703) 836-6620
;       TELEFAX:  (703) 836-2021
;   INFORMATION FOR SEQ ID NO:  2:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  29 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  double
;       TOPOLOGY:  circular
;     MOLECULE TYPE:  other nucleic acid
;       DESCRIPTION:  /desc = "portion of plasmid pTG150"
;     FEATURE:
;       NAME/KEY:  CDS
;       LOCATION:  15..29
;     FEATURE:
;       NAME/KEY:  misc_feature
;       LOCATION:  15..29
;       OTHER INFORMATION:  /note= "5'end of rabies
;       OTHER INFORMATION:  glycoprotein cDNA inserted into plasmid pTG150."
;     FEATURE:
;       NAME/KEY:  mat_peptide
;       LOCATION:  27..29
US-08-480-736-2

  Query Match             64.0%;  Score 12.8;  DB 2;  Length 29;
  Best Local Similarity   87.5%;  Pred. No. 6.5e+02;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       2 CTGAGAAACCTTCTTT 17
           ||||| |||| |||||
Db      26 CTGAGGAACCATCTTT 11


Search completed: November 23, 2002, 07:07:31
Job time : 23.3 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 06:42:25 ; Search time 16.8 Seconds
(without alignments)
450.869 Million cell updates/sec

Title: US-09-296-264-21
Perfect score: 20
Sequence: 1 gctgagaaaccttcttttgc 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 335578 seqs, 189365133 residues

Total number of hits satisfying chosen parameters: 195614

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	13.2	66.0	81	10	US-09-878-574-235	Sequence 235, App
c	2	13.2	66.0	90	10	US-09-864-761-27537	Sequence 27537, A
c	3	12.8	64.0	87	10	US-09-864-761-24810	Sequence 24810, A
	4	12.8	64.0	95	10	US-09-815-242-1296	Sequence 1296, Ap
	5	12.6	63.0	61	10	US-09-865-807-60	Sequence 60, Appl
c	6	12.6	63.0	80	10	US-09-864-761-18877	Sequence 18877, A
c	7	12.2	61.0	23	10	US-09-216-393-227	Sequence 227, App
c	8	12.2	61.0	26	9	US-10-097-556-1	Sequence 1, Appli
	9	12.2	61.0	29	10	US-09-745-763-156	Sequence 156, App
c	10	12.2	61.0	31	10	US-09-801-274-601	Sequence 601, App
c	11	12.2	61.0	31	10	US-09-801-274-1711	Sequence 1711, Ap
	12	12.2	61.0	34	10	US-09-732-561-1	Sequence 1, Appli
c	13	12.2	61.0	66	10	US-09-783-590-1423	Sequence 1423, Ap
	14	12.2	61.0	69	10	US-09-797-207-11	Sequence 11, Appl
	15	12.2	61.0	83	10	US-09-923-876-1045	Sequence 1045, Ap
c	16	12.2	61.0	96	10	US-09-923-876-4080	Sequence 4080, Ap
c	17	12	60.0	24	9	US-09-905-291A-20	Sequence 20, Appl
c	18	12	60.0	24	10	US-09-909-320-20	Sequence 20, Appl
c	19	12	60.0	24	10	US-09-909-088B-20	Sequence 20, Appl
	20	12	60.0	25	10	US-09-804-661-4	Sequence 4, Appli
c	21	12	60.0	82	10	US-09-864-761-31287	Sequence 31287, A
c	22	12	60.0	94	10	US-09-923-876-1103	Sequence 1103, Ap
c	23	12	60.0	95	10	US-09-969-373-513	Sequence 513, App
c	24	11.8	59.0	20	10	US-09-752-983-111	Sequence 111, App
	25	11.8	59.0	20	10	US-09-969-373-2375	Sequence 2375, Ap
	26	11.8	59.0	20	10	US-09-969-373-2564	Sequence 2564, Ap
c	27	11.8	59.0	21	12	US-10-099-355-6	Sequence 6, Appli
	28	11.8	59.0	28	10	US-09-893-238-78	Sequence 78, Appl
c	29	11.8	59.0	29	9	US-09-949-134A-1	Sequence 1, Appli
c	30	11.8	59.0	31	12	US-10-007-805-562	Sequence 562, App
	31	11.8	59.0	68	9	US-09-949-134A-25	Sequence 25, Appl
c	32	11.6	58.0	31	10	US-09-801-274-461	Sequence 461, App
c	33	11.6	58.0	31	10	US-09-801-274-1396	Sequence 1396, Ap
	34	11.6	58.0	36	10	US-09-968-355-8	Sequence 8, Appli
	35	11.6	58.0	81	10	US-09-864-761-18385	Sequence 18385, A
c	36	11.6	58.0	94	10	US-09-294-093B-4414	Sequence 4414, Ap
	37	11.6	58.0	96	8	US-08-896-322-2	Sequence 2, Appli
c	38	11.6	58.0	98	10	US-09-864-761-29457	Sequence 29457, A
c	39	11.6	58.0	99	10	US-09-864-761-23701	Sequence 23701, A
c	40	11.4	57.0	36	10	US-09-935-727-42	Sequence 42, Appl
c	41	11.4	57.0	47	9	US-09-853-526-245	Sequence 245, App
c	42	11.4	57.0	47	9	US-09-853-526-246	Sequence 246, App
c	43	11.4	57.0	47	9	US-09-853-526-322	Sequence 322, App
c	44	11.4	57.0	47	9	US-09-853-526-323	Sequence 323, App
c	45	11.4	57.0	47	10	US-09-901-484A-245	Sequence 245, App

ALIGNMENTS

RESULT 1
US-09-878-574-235
; Sequence 235, Application US/09878574
; Patent No. US20020110548A1
; GENERAL INFORMATION:
; APPLICANT: Byrum, Joseph R.
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Thompson, Michael D.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(15401)B
; CURRENT APPLICATION NUMBER: US/09/878,574
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 09/333,535
; PRIOR FILING DATE: 1999-06-14
; NUMBER OF SEQ ID NOS: 15775
; SEQ ID NO 235
; LENGTH: 81
; TYPE: DNA
; ORGANISM: Glycine max
; OTHER INFORMATION: Clone ID: LIB3028-053-Q1-B1-E6
US-09-878-574-235

Query Match 66.0%; Score 13.2; DB 10; Length 81;
Best Local Similarity 83.3%; Pred. No. 1e+03;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy       2 CTGAGAAACCTTCTTTTG 19
           |||  ||||||| |||||
Db      53 CTGGTAAACCTTGTTTTG 70
```

RESULT 2
US-09-864-761-27537/c
; Sequence 27537, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO

```
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 27537
;   LENGTH: 90
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AC005723.1
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.61
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.64
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.69
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.63
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.49
;   OTHER INFORMATION: NT HIT: U22463.1, EVALUE 7.40e-01
;   OTHER INFORMATION: EST_HUMAN HIT: AA404576.1, EVALUE 2.80e-01
;   OTHER INFORMATION: SWISSPROT HIT: P49581, EVALUE 7.30e+00
US-09-864-761-27537

  Query Match             66.0%;  Score 13.2;  DB 10;  Length 90;
  Best Local Similarity   83.3%;  Pred. No. 1e+03;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       3 TGAGAAACCTTCTTTTGC 20
           ||| || ||| |||||||
Db      76 TGACAATCCTCCTTTTGC 59


RESULT 3 .
US-09-864-761-24810/c
; Sequence 24810, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 24810
;   LENGTH: 87
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AL158148.1
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 7.3
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL =   5
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 5.9
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4.6
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 7.5
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 7.4
;   OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 6.6
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL =   6
;   OTHER INFORMATION: NT HIT: U38906.1, EVALUE 1.80e-01
US-09-864-761-24810

  Query Match             64.0%;  Score 12.8;  DB 10;  Length 87;
  Best Local Similarity   87.5%;  Pred. No. 1.6e+03;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 GCTGAGAAACCTTCTT 16
           | |||||||||  ||||
Db      48 GGTGAGAAACCGTCTT 33


RESULT 4
```

US-09-815-242-1296
; Sequence 1296, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1296
; LENGTH: 95
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-815-242-1296

Query Match 64.0%; Score 12.8; DB 10; Length 95;
Best Local Similarity 87.5%; Pred. No. 1.6e+03;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GCTGAGAAACCTTCTT 16
|| ||||||| |||||
Db 50 GCGGAGAAACTTTCTT 65

RESULT 5
US-09-865-807-60
; Sequence 60, Application US/09865807
; Patent No. US20020068334A1
; GENERAL INFORMATION:
; APPLICANT: Carrino, John J.
; APPLICANT: Gerrue, Louis O.
; APPLICANT: Diver, Jonathan M.
; TITLE OF INVENTION: Multiplex Amplification and Separation of Nucleic Acid
; TITLE OF INVENTION: Sequences Using Ligation-Dependant Strand Displacement
; TITLE OF INVENTION: Amplification and Bioelectronic Chip Technology
; FILE REFERENCE: 265/018 Nanogen
; CURRENT APPLICATION NUMBER: US/09/865,807
; CURRENT FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 60
; LENGTH: 61
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Ligation probe
US-09-865-807-60

Query Match 63.0%; Score 12.6; DB 10; Length 61;
Best Local Similarity 78.9%; Pred. No. 1.9e+03;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 2 CTGAGAAACCTTCTTTTGC 20
|||||| || ||||| ||
Db 29 CTGAGATCCCCTCTTTGGC 47

RESULT 6
US-09-864-761-18877/c
; Sequence 18877, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 18877
; LENGTH: 80
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL109615.8
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 3.8
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 9.5
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 5.3
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.1

```
;   OTHER INFORMATION: EST_HUMAN HIT: AA932558.1, EVALUE 1.00e-37
;   OTHER INFORMATION: SWISSPROT HIT: P34492, EVALUE 9.30e+00
US-09-864-761-18877

  Query Match             63.0%;  Score 12.6;  DB 10;  Length 80;
  Best Local Similarity   78.9%;  Pred. No. 2e+03;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      2 CTGAGAAACCTTCTTTTGC 20
          ||||||| |   |||||||
Db     19 CTGAGAACCACCCTTTTGC 1


RESULT 7
US-09-216-393-227/c
; Sequence 227, Application US/09216393
; Patent No. US20010014447A1
; GENERAL INFORMATION:
;  APPLICANT: Milhausen, Michael James
;  TITLE OF INVENTION: TOXOPLASMA GONDII PROTEINS, NUCLEIC ACID MOLECULES, AND
;  TITLE OF INVENTION: USES THEREOF
;  FILE REFERENCE: TX-1-C2
;  CURRENT APPLICATION NUMBER: US/09/216,393
;  CURRENT FILING DATE: 1998-12-18
;  EARLIER APPLICATION NUMBER: 08/994,825
;  EARLIER FILING DATE: 1997-12-19
;  NUMBER OF SEQ ID NOS: 364
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 227
;   LENGTH: 23
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:  Synthetic
;   OTHER INFORMATION: Primer
US-09-216-393-227

  Query Match             61.0%;  Score 12.2;  DB 10;  Length 23;
  Best Local Similarity   82.4%;  Pred. No. 2.4e+03;
  Matches   14;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      4 GAGAAACCTTCTTTTGC 20
          |||| |||| ||||| |
Db     23 GAGACACCTCCTTTTCC 7


RESULT 8
US-10-097-556-1/c
; Sequence 1, Application US/10097556
; Patent No. US20020168760A1
; GENERAL INFORMATION:
;  APPLICANT: Dornburg, Ralph
;  APPLICANT:  Schnell, Matthias J.
;  APPLICANT:  Dietzschold, Bernhard
;  TITLE OF INVENTION: Retroviral Vectors for Gene Transfer
;  TITLE OF INVENTION:  into Neuronal Cells
;  FILE REFERENCE: DOR01.NP004
;  CURRENT APPLICATION NUMBER: US/10/097,556
;  CURRENT FILING DATE:  2002-03-13
;  NUMBER OF SEQ ID NOS: 2
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
;   LENGTH: 26
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: PCR primers
US-10-097-556-1

  Query Match             61.0%;  Score 12.2;  DB 9;  Length 26;
  Best Local Similarity   82.4%;  Pred. No. 2.5e+03;
  Matches   14;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 CTGAGAAACCTTCTTTT 18
          ||||| |||| |||| |
Db     22 CTGAGGAACCATCTTCT 6


RESULT 9
US-09-745-763-156
; Sequence 156, Application US/09745763
; Patent No. US20020065394A1
;  GENERAL INFORMATION:
;       APPLICANT: Jacobs, Kenneth
;                  McCoy, John M.
;                  LaVallie, Edward R.
;                  Collins-Racie, Lisa A.
;                  Evans, Cheryl
;                  Merberg, David
;                  Treacy, Maurice
;                  Spaulding, Vikki
;       TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
;                           ENCODING THEM
;       NUMBER OF SEQUENCES: 219
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Genetics Institute, Inc.
;            STREET: 87 CambridgePark Drive
;            CITY: Cambridge
;            STATE: MA
;            COUNTRY: U.S.A.
;            ZIP: 02140
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/745,763
;            FILING DATE: 18-Jun-2000
;            CLASSIFICATION: <Unknown>
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Sprunger, Suzanne A.
;            REGISTRATION NUMBER: 41,323
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (617) 498-8284
;            TELEFAX: (617) 876-5851
;   INFORMATION FOR SEQ ID NO: 156:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 29 base pairs
;            TYPE: nucleic acid
;            STRANDEDNESS: single
;            TOPOLOGY: linear
;       MOLECULE TYPE: other nucleic acid
;            DESCRIPTION: /desc = "oligonucleotide"
;       SEQUENCE DESCRIPTION: SEQ ID NO: 156:
US-09-745-763-156

  Query Match             61.0%;  Score 12.2;  DB 10;  Length 29;
  Best Local Similarity   82.4%;  Pred. No. 2.5e+03;
  Matches   14;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 CTGAGAAACCTTCTTTT 18
          || |||| ||||| |||
Db      3 CTTAGAAGCCTTCATTT 19


RESULT 10
US-09-801-274-601/c
; Sequence 601, Application US/09801274
; Patent No. US20020032319A1
; GENERAL INFORMATION:
;  APPLICANT: Cargill, Michele
;  APPLICANT:  Ireland, James S.
;  APPLICANT:  Lander, Eric S.
;  TITLE OF INVENTION: HUMAN SINGLE NUCLEOTIDE POLYMORPHISMS
;  FILE REFERENCE: 2825.2009-001
```

```
;  CURRENT APPLICATION NUMBER: US/09/801,274
;  CURRENT FILING DATE:  2001-03-07
;  PRIOR APPLICATION NUMBER: US 60/187,510
;  PRIOR FILING DATE: 2000-03-07
;  PRIOR APPLICATION NUMBER: US 60/206,129
;  PRIOR FILING DATE: 2000-05-22
;  NUMBER OF SEQ ID NOS: 1802
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 601
;   LENGTH: 31
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-801-274-601

  Query Match             61.0%;  Score 12.2;  DB 10;  Length 31;
  Best Local Similarity   73.7%;  Pred. No. 2.6e+03;
  Matches   14;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 CTGAGAAACCTTCTTTTGC 20
           || ||||  ||: ||||||
Db      27 CTTAGAATGCTWTTTTTGC 9


RESULT 11
US-09-801-274-1711/c
; Sequence 1711, Application US/09801274
; Patent No. US20020032319A1
; GENERAL INFORMATION:
;  APPLICANT: Cargill, Michele
;  APPLICANT:  Ireland, James S.
;  APPLICANT:  Lander, Eric S.
;  TITLE OF INVENTION: HUMAN SINGLE NUCLEOTIDE POLYMORPHISMS
;  FILE REFERENCE: 2825.2009-001
;  CURRENT APPLICATION NUMBER: US/09/801,274
;  CURRENT FILING DATE:  2001-03-07
;  PRIOR APPLICATION NUMBER: US 60/187,510
;  PRIOR FILING DATE: 2000-03-07
;  PRIOR APPLICATION NUMBER: US 60/206,129
;  PRIOR FILING DATE: 2000-05-22
;  NUMBER OF SEQ ID NOS: 1802
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1711
;   LENGTH: 31
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-801-274-1711

  Query Match             61.0%;  Score 12.2;  DB 10;  Length 31;
  Best Local Similarity   73.7%;  Pred. No. 2.6e+03;
  Matches   14;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 GCTGAGAAACCTTCTTTTG 19
           ||||||:||   ||| |||
Db      22 GCTGAGRAATGGTCTGTTG 4


RESULT 12
US-09-732-561-1
; Sequence 1, Application US/09732561
; Patent No. US20020035738A1
;  GENERAL INFORMATION:
;    APPLICANT:  Thomma, Bart
;    APPLICANT:  Terras, Franky
;    APPLICANT:  Penninckx, Iris
;    APPLICANT:  Manners, John
;    APPLICANT:  Kazan, Kemal
;    APPLICANT:  Broekaert, Willem
;    TITLE OF INVENTION:  Plant Protection Method
;    NUMBER OF SEQUENCES:  24
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  ZENECA Ag Products
;      STREET:  1800 Concord Pike
;      CITY:  Wilmington
;      STATE:  DE
;      COUNTRY:  USA
;      ZIP:  19850
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/732,561
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  09/202,638
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/GB97/01672
;      FILING DATE:  20-JUN-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hohenschutz, Liza D.
;      REGISTRATION NUMBER:  33,712
;      REFERENCE/DOCKET NUMBER:  PPD 50165/UST
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (302) 886-1699
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  34 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    IMMEDIATE SOURCE:
;      CLONE:  OWB260 PRIMER
US-09-732-561-1

  Query Match             61.0%;  Score 12.2;  DB 10;  Length 34;
  Best Local Similarity   82.4%;  Pred. No. 2.6e+03;
  Matches   14;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       4 GAGAAACCTTCTTTTGC 20
           |||||| ||| || |||
Db       3 GAGAAAGCTTGTTGTGC 19


RESULT 13
US-09-783-590-1423/c
; Sequence 1423, Application US/09783590
; Patent No. US20020110850A1
; GENERAL INFORMATION:
;  APPLICANT: Dillon, Patrick J.
;  APPLICANT:  Haseltine, William A.
;  APPLICANT:  Li, Haodong
;  APPLICANT:  Rosen, Craig A.
;  APPLICANT:  Ruben, Steven M.
;  TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 16.2
;  FILE REFERENCE: PO-16.2C1
;  CURRENT APPLICATION NUMBER: US/09/783,590
;  CURRENT FILING DATE:  2000-02-15
;  PRIOR APPLICATION NUMBER: 08/420,856
;  PRIOR FILING DATE: 1995-04-12
;  PRIOR APPLICATION NUMBER: 08/346,731
;  PRIOR FILING DATE: 1994-11-21
;  NUMBER OF SEQ ID NOS: 12485
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1423
;   LENGTH: 66
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc feature
;   LOCATION: (8)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc feature
```

```
;   LOCATION: (41)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc feature
;   LOCATION: (55)
;   OTHER INFORMATION: n equals a,t,g, or c
US-09-783-590-1423

  Query Match             61.0%;  Score 12.2;  DB 10;  Length 66;
  Best Local Similarity   77.8%;  Pred. No. 3e+03;
  Matches   14;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      3 TGAGAAACCTTCTTTTGC 20
          |||| | | |||||| ||
Db     23 TGAGCAGCTTTCTTTNGC 6


RESULT 14
US-09-797-207-11
; Sequence 11, Application US/09797207
; Patent No. US20020098563A1
; GENERAL INFORMATION:
;  APPLICANT: KORCZAK, BOZENA
;  TITLE OF INVENTION: NOVEL CORE 2 BETA-1, 6-N-ACETYLGLYCOSAMINYLTRANSFERASE
;  TITLE OF INVENTION: GENE
;  FILE REFERENCE: GLYCO-7P1
;  CURRENT APPLICATION NUMBER: US/09/797,207
;  CURRENT FILING DATE: 2001-03-02
;  EARLIER APPLICATION NUMBER: 09/495,913
;  EARLIER FILING DATE: 2000-02-02
;  EARLIER APPLICATION NUMBER: 60/118,674
;  EARLIER FILING DATE: 1999-02-03
;  NUMBER OF SEQ ID NOS: 20
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
;   LENGTH: 69
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Recombinant
;   OTHER INFORMATION: DNA
US-09-797-207-11

  Query Match             61.0%;  Score 12.2;  DB 10;  Length 69;
  Best Local Similarity   82.4%;  Pred. No. 3e+03;
  Matches   14;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 GCTGAGAAACCTTCTTT 17
          |||  ||||| ||||||
Db     49 GCTCTGAAACTTTCTTT 65


RESULT 15
US-09-923-876-1045
; Sequence 1045, Application US/09923876
; Patent No. US20020013958A1
; GENERAL INFORMATION:
;  APPLICANT: Lalgudi, Raghunath V.
;  APPLICANT:  Kamigaki, Laura Y. (Ito)
;  APPLICANT:  Sherman, Bradley K.
;  TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN SEEDLING
;  FILE REFERENCE: PL-0012-1 CON
;  CURRENT APPLICATION NUMBER: US/09/923,876
;  CURRENT FILING DATE:  2001-08-06
;  PRIOR APPLICATION NUMBER: 09/298,329
;  PRIOR FILING DATE: 1999-04-21
;  PRIOR APPLICATION NUMBER: 60/085,331
;  PRIOR FILING DATE: 1998-05-05
;  NUMBER OF SEQ ID NOS: 6332
;  SOFTWARE: PERL Program
; SEQ ID NO 1045
;   LENGTH: 83
;   TYPE: DNA
;   ORGANISM: Zea mays
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. US20020013958A1 700158023H1
;   NAME/KEY: unsure
;   LOCATION: 38
;   OTHER INFORMATION: a, t, c, g, or other
US-09-923-876-1045

  Query Match             61.0%;  Score 12.2;  DB 10;  Length 83;
  Best Local Similarity   82.4%;  Pred. No. 3.1e+03;
  Matches   14;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      4 GAGAAACCTTCTTTTGC 20
          ||||||  | |||||||
Db     66 GAGAAAATTGCTTTTGC 82


Search completed: November 23, 2002, 07:10:38
Job time : 16.8 secs
```

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 26, 2002, 04:16:10 ; Search time 798.5 Seconds
(without alignments)
405.648 Million cell updates/sec

Title: US-09-296-264-21
Perfect score: 20
Sequence: 1 gctgagaaaccttcttttgc 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 357874

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	15.2	76.0	50	9	AU102562	AU102562 AU102562
c	2	14.4	72.0	39	17	AZ307139	AZ307139 1M0008G03
c	3	14.4	72.0	65	14	BQ243021	BQ243021 TaE15020C
c	4	14.2	71.0	23	17	BH853341	BH853341 SALK_0767
	5	14.2	71.0	93	17	BH846776	BH846776 SALK_0103
	6	14.2	71.0	94	17	BH846812	BH846812 SALK_0105
	7	13.8	69.0	92	13	BJ081265	BJ081265 BJ081265
	8	13.6	68.0	40	9	AA725970	AA725970 vu83g12.r
c	9	13.6	68.0	50	9	AU102560	AU102560 AU102560
c	10	13.6	68.0	50	9	AU102565	AU102565 AU102565
c	11	13.6	68.0	76	17	B35044	B35044 HS-1026-B1-
c	12	13.4	67.0	66	14	BQ568492	BQ568492 gi109b06.
c	13	13.4	67.0	76	9	AA512483	AA512483 vj16h04.r
c	14	13.2	66.0	64	17	CNS03DGG	AL239065 Tetraodon
c	15	13.2	66.0	72	9	AA103844	AA103844 mo44c05.r
	16	13.2	66.0	79	9	AA515945	AA515945 ni30e06.s
	17	13.2	66.0	82	17	AL766061	AL766061 Arabidops
	18	13.2	66.0	87	17	AL766060	AL766060 Arabidops
c	19	13.2	66.0	88	17	AZ788933	AZ788933 2M0036D21
	20	13.2	66.0	92	17	AZ799203	AZ799203 2M0056L13
c	21	13.2	66.0	93	9	AI222301	AI222301 qg97e05.x
c	22	13.2	66.0	93	17	AZ575546	AZ575546 AST-T23F0
c	23	13.2	66.0	95	9	AA990332	AA990332 ua64a05.r
c	24	13.2	66.0	96	14	H55422	H55422 CHR220361 C
	25	13.2	66.0	96	17	AZ795342	AZ795342 2M0049G02
c	26	13.2	66.0	100	10	AW698411	AW698411 G259 glan
	27	12.8	64.0	26	17	AZ602086	AZ602086 1M0420K07
	28	12.8	64.0	36	13	BJ054767	BJ054767 BJ054767
	29	12.8	64.0	49	9	AA454680	AA454680 zx76b01.s
	30	12.8	64.0	75	9	AI540552	AI540552 tq36d03.x
c	31	12.8	64.0	88	9	AA946725	AA946725 of13h02.s
c	32	12.8	64.0	91	10	BE546379	BE546379 601071061
c	33	12.8	64.0	94	9	AU254804	AU254804 AU254804
c	34	12.8	64.0	94	9	AA434287	AA434287 zw30f03.r
	35	12.8	64.0	94	12	BG315742	BG315742 sab84f02.
	36	12.6	63.0	39	17	AZ772153	AZ772153 1M0574N16
c	37	12.6	63.0	43	9	AI362608	AI362608 qv72h02.x
	38	12.6	63.0	50	9	AU104160	AU104160 AU104160
c	39	12.6	63.0	50	9	AU107694	AU107694 AU107694
	40	12.6	63.0	73	17	AZ419596	AZ419596 1M0196K12
c	41	12.6	63.0	76	17	BH865009	BH865009 SALK_0972
c	42	12.6	63.0	78	17	B35501	B35501 HS-1029-A1-
c	43	12.6	63.0	79	9	AI095995	AI095995 SWOvL3CAN
c	44	12.6	63.0	82	9	AA458985	AA458985 zx88h08.s
	45	12.6	63.0	83	17	AZ337749	AZ337749 1M0068K01

ALIGNMENTS

RESULT 1
AU102562/c
LOCUS AU102562 50 bp mRNA linear EST 30-AUG-2001
DEFINITION AU102562 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone COL09872, mRNA sequence.
ACCESSION AU102562
VERSION AU102562.1 GI:13552083
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 50)
AUTHORS Suzuki,Y., Taira,H., Tsunoda,T., Mizushima-Sugano,J., Sese,J., Hata,H., Ota,T., Isogai,T., Tanaka,T., Morishita,S., Okubo,K., Sakaki,Y., Nakamura,Y., Suyama,A. and Sugano,S.
TITLE Diverse transcriptional initiation revealed by fine, large-scale mapping of mRNA start sites
JOURNAL EMBO Rep. 2 (5), 388-393 (2001)
MEDLINE 21270072
COMMENT Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: ysuzuki@ims.u-tokyo.ac.jp
Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano,S. Construction and characterization of a full length-enriched and a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).
FEATURES Location/Qualifiers

```
     source          1. .50
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="COL09872"
                     /clone_lib="Sugano Homo sapiens cDNA library"
                     /note="Differential display comparison of untreated and
                     dimethylfumarate treated U937 cells"
BASE COUNT        9 a     14 c     16 g     11 t
ORIGIN

  Query Match             76.0%;  Score 15.2;  DB 9;  Length 50;
  Best Local Similarity   85.0%;  Pred. No. 2.4e+03;
  Matches   17;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 GCTGAGAAACCTTCTTTTGC 20
          |||||| |||| ||| ||||
Db     21 GCTGAGGAACCGTCTATTGC 2


RESULT 2
AZ307139/c
LOCUS       AZ307139                 39 bp    DNA     linear   GSS 29-SEP-2000
DEFINITION  1M0008G03R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0008G03 R, DNA sequence.
ACCESSION   AZ307139
VERSION     AZ307139.1  GI:10345842
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 39)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0008  row: G  column: 03
            Seq primer: CACACAGGAAACAGCTATGACC
            Class: plasmid ends
            High quality sequence stop: 39.
FEATURES             Location/Qualifiers
     source          1. .39
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC1M0008G03"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
BASE COUNT        6 a      9 c     10 g     14 t
ORIGIN

  Query Match             72.0%;  Score 14.4;  DB 17;  Length 39;
  Best Local Similarity   93.8%;  Pred. No. 5.8e+03;
  Matches   15;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 GCTGAGAAACCTTCTT 16
          |||||||||||| ||||
Db     16 GCTGAGAAACCCTCTT 1


RESULT 3
BQ243021/c
LOCUS       BQ243021                 65 bp    mRNA    linear   EST 03-MAY-2002
DEFINITION  TaE15020C08F TaE15 Triticum aestivum cDNA clone TaE15020C08F, mRNA
            sequence.
ACCESSION   BQ243021
VERSION     BQ243021.1  GI:20438884
KEYWORDS    EST.
SOURCE      bread wheat.
  ORGANISM  Triticum aestivum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
            ; Triticeae; Triticum.
REFERENCE   1  (bases 1 to 65)
  AUTHORS   Cloutier,S.
  TITLE     Wheat functional genomics - Glenlea developing seeds cDNA libraries
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Dr. Sylvie Cloutier
            Cereal Research Centre, Agriculture and Agri-food Canada
            195 Dafoe Rd, Winnipeg, MB, Canada R3T 2M9
            Tel: (204) 983-2340
            Fax: (204) 983-4604
            Email: scloutier@em.agr.ca
            was cloned directionally, not all sequences generated with reverse
            primer were from the 5' end (same with forward primer and 3' end).
            Average insert size is >1.4 kb
            Plate: 020  row: C  column: 08
            Seq primer: M13 Forward.
FEATURES             Location/Qualifiers
     source          1. .65
                     /organism="Triticum aestivum"
                     /cultivar="Glenlea"
                     /db_xref="taxon:4565"
                     /clone="TaE15020C08F"
                     /clone_lib="TaE15"
                     /tissue_type="developing seeds"
                     /dev_stage="15 days after anthesis"
                     /lab_host="E. coli DH10B"
                     /note="Vector: pCMV-SPORT6.0 (Invitrogen Technologies);
                     Site_1: NotI; Site_2: MluI; mRNA obtained from wheat seeds
                     of cultivar Glenlea 15 days post-anthesis"
BASE COUNT       22 a     12 c     10 g     21 t
ORIGIN

  Query Match             72.0%;  Score 14.4;  DB 14;  Length 65;
  Best Local Similarity   93.8%;  Pred. No. 6.5e+03;
  Matches   15;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      5 AGAAACCTTCTTTTGC 20
          ||||||||| ||||||
Db     65 AGAAACCTTTTTTTGC 50


RESULT 4
```

```
BH853341/c
LOCUS       BH853341                 23 bp    DNA     linear   GSS 13-JUN-2002
DEFINITION  SALK_076754.44.75.x Arabidopsis thaliana TDNA insertion lines
            Arabidopsis thaliana genomic clone SALK_076754.44.75.x, DNA
            sequence.
ACCESSION   BH853341
VERSION     BH853341.1  GI:21424212
KEYWORDS    GSS.
SOURCE      thale cress.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 23)
  AUTHORS   Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R., Gadrinab
            ,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L., Shinn,P.
            , Zimmerman,J. and Ecker,J.R.
  TITLE     A Sequence-Indexed Library of Insertion Mutations in the
            Arabidopsis Genome
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Joseph R. Ecker
            Salk Institute Genomic Analysis Laboratory (SIGnAL)
            The Salk Institute for Biological Studies
            10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
            Tel: 858 453 4100 x1752
            Fax: 858 558 6379
            Email: ecker@salk.edu
            This is single pass sequence recovered from the left border of
            TDNA. This sequence lies within an annotated intron of At4g31430.
            Class: TDNA tagged.
FEATURES             Location/Qualifiers
     source          1. .23
                     /organism="Arabidopsis thaliana"
                     /strain="Columbia 0"
                     /db_xref="taxon:3702"
                     /clone="SALK_076754.44.75.x"
                     /clone_lib="Arabidopsis thaliana TDNA insertion lines"
                     /note="PCR was performed on Arabidopsis thaliana lines
                     each of which contains  one or more TDNA insertion
                     elements. The resultant fragment for each  line was
                     directly sequenced to determine the genomic sequence at
                     the  site of insertion. Details of the protocols used can
                     be found at  http://signal.salk.edu/tdna_protocols.html"
BASE COUNT        8 a      2 c      6 g      7 t
ORIGIN

  Query Match             71.0%;  Score 14.2;  DB 17;  Length 23;
  Best Local Similarity   84.2%;  Pred. No. 6.6e+03;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 CTGAGAAACCTTCTTTTGC 20
           ||||  ||||| |||||||
Db      19 CTGAACAACCTACTTTTGC 1


RESULT 5
BH846776
LOCUS       BH846776                 93 bp    DNA     linear   GSS 13-JUN-2002
DEFINITION  SALK_010338.36.10.x Arabidopsis thaliana TDNA insertion lines
            Arabidopsis thaliana genomic clone SALK_010338.36.10.x, DNA
            sequence.
ACCESSION   BH846776
VERSION     BH846776.1  GI:21417647
KEYWORDS    GSS.
SOURCE      thale cress.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 93)
  AUTHORS   Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R., Gadrinab
            ,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L., Shinn,P.
            , Zimmerman,J. and Ecker,J.R.
  TITLE     A Sequence-Indexed Library of Insertion Mutations in the
            Arabidopsis Genome
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Joseph R. Ecker
            Salk Institute Genomic Analysis Laboratory (SIGnAL)
            The Salk Institute for Biological Studies
            10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
            Tel: 858 453 4100 x1752
            Fax: 858 558 6379
            Email: ecker@salk.edu
            This is single pass sequence recovered from the left border of
            TDNA. This sequence lies within an annotated exon of At3g55190.
            Class: TDNA tagged.
FEATURES             Location/Qualifiers
     source          1. .93
                     /organism="Arabidopsis thaliana"
                     /strain="Columbia 0"
                     /db_xref="taxon:3702"
                     /clone="SALK_010338.36.10.x"
                     /clone_lib="Arabidopsis thaliana TDNA insertion lines"
                     /note="PCR was performed on Arabidopsis thaliana lines
                     each of which contains  one or more TDNA insertion
                     elements. The resultant fragment for each  line was
                     directly sequenced to determine the genomic sequence at
                     the  site of insertion. Details of the protocols used can
                     be found at  http://signal.salk.edu/tdna_protocols.html"
BASE COUNT       22 a     25 c     11 g     29 t      6 others
ORIGIN

  Query Match             71.0%;  Score 14.2;  DB 17;  Length 93;
  Best Local Similarity   84.2%;  Pred. No. 8.9e+03;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 CTGAGAAACCTTCTTTTGC 20
           | |  ||||||||||||||
Db      33 CAGCAAAACCTTCTTTTGC 51


RESULT 6
BH846812
LOCUS       BH846812                 94 bp    DNA     linear   GSS 13-JUN-2002
DEFINITION  SALK_010553.40.80.x Arabidopsis thaliana TDNA insertion lines
            Arabidopsis thaliana genomic clone SALK_010553.40.80.x, DNA
            sequence.
ACCESSION   BH846812
VERSION     BH846812.1  GI:21417683
KEYWORDS    GSS.
SOURCE      thale cress.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 94)
  AUTHORS   Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R., Gadrinab
            ,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L., Shinn,P.
            , Zimmerman,J. and Ecker,J.R.
  TITLE     A Sequence-Indexed Library of Insertion Mutations in the
            Arabidopsis Genome
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Joseph R. Ecker
            Salk Institute Genomic Analysis Laboratory (SIGnAL)
            The Salk Institute for Biological Studies
            10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
            Tel: 858 453 4100 x1752
            Fax: 858 558 6379
            Email: ecker@salk.edu
            This is single pass sequence recovered from the left border of
            TDNA. This sequence lies within an annotated exon of At3g55190.
            Class: TDNA tagged.
FEATURES             Location/Qualifiers
     source          1. .94
                     /organism="Arabidopsis thaliana"
                     /strain="Columbia 0"
```

```
                     /db_xref="taxon:3702"
                     /clone="SALK_010553.40.80.x"
                     /clone_lib="Arabidopsis thaliana TDNA insertion lines"
                     /note="PCR was performed on Arabidopsis thaliana lines
                     each of which contains  one or more TDNA insertion
                     elements. The resultant fragment for each  line was
                     directly sequenced to determine the genomic sequence at
                     the  site of insertion. Details of the protocols used can
                     be found at  http://signal.salk.edu/tdna_protocols.html"
BASE COUNT       27 a     25 c     12 g     30 t
ORIGIN

  Query Match             71.0%;  Score 14.2;  DB 17;  Length 94;
  Best Local Similarity   84.2%;  Pred. No. 8.9e+03;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 CTGAGAAACCTTCTTTTGC 20
          | |  ||||||||||||||||
Db     34 CAGCAAAACCTTCTTTTGC 52


RESULT 7
BJ081265
LOCUS       BJ081265                 92 bp    mRNA    linear   EST 12-DEC-2001
DEFINITION  BJ081265 NIBB Mochii normalized Xenopus tailbud library Xenopus
            laevis cDNA clone XL071b22 3', mRNA sequence.
ACCESSION   BJ081265
VERSION     BJ081265.1  GI:17575358
KEYWORDS    EST.
SOURCE      African clawed frog.
  ORGANISM  Xenopus laevis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus.
REFERENCE   1  (bases 1 to 92)
  AUTHORS   Kitayama,A., Terasaka,C., Mochii,M., Ueno,N., Shin-i,T. and Kohara
            ,Y.
  TITLE     Expressed genes in X. laevis embryo
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Tadasu Shin-i
            Center For Genetic Resource Information
            National Institute of Genetics
            1111 Yata, Mishima, Shizuoka 411-8540, Japan
            Tel: 81-559-81-6856
            Fax: 81-559-81-6855
            Email: tshini@genes.nig.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .92
                     /organism="Xenopus laevis"
                     /db_xref="taxon:8355"
                     /clone="XL071b22"
                     /clone_lib="NIBB Mochii normalized Xenopus tailbud
                     library"
                     /tissue_type="whole embryo"
                     /dev_stage="stage 25"
                     /note="Vector: pBSRN3; Site_1: NotI; Site_2: EcoRI; cDNAs
                     were oligo-dT primed and directionally cloned. Staging
                     according to  Nieuwkoop and Faber. Library is subtracted
                     and was constructed by N. Garrett  and A.M. Zorn,
                     (Wellcome/CRC Institute). "
BASE COUNT       23 a     21 c     13 g     30 t      5 others
ORIGIN

  Query Match             69.0%;  Score 13.8;  DB 13;  Length 92;
  Best Local Similarity   88.2%;  Pred. No. 1.4e+04;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      2 CTGAGAAACCTTCTTTT 18
          |||  |||||||| ||||
Db     41 CTGGGAAACCTTTTTTT 57


RESULT 8
AA725970
LOCUS       AA725970                 40 bp    mRNA    linear   EST 02-JAN-1998
DEFINITION  vu83g12.r1 Stratagene mouse skin (#937313) Mus musculus cDNA clone
            IMAGE:1198054 5' similar to TR:Q28576 Q28576 HAIR KERATIN CYSTEINE
            RICH PROTEIN ;, mRNA sequence.
ACCESSION   AA725970
VERSION     AA725970.1  GI:2743677
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 40)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:645150
            Trace considered overall poor quality
            Possible reversed clone: similarity on wrong strand
            Seq primer: -28m13 rev1 ET from Amersham
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .40
                     /organism="Mus musculus"
                     /strain="C57BL/6"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:1198054"
                     /clone_lib="Stratagene mouse skin (#937313)"
                     /sex="females"
                     /tissue_type="whole skin"
                     /dev_stage="11 weeks old"
                     /lab_host="SOLR (kanamycin resistant)"
                     /note="Organ: skin; Vector: pBluescript SK-; Site_1: EcoRI
                     ; Site_2: XhoI; Cloned unidirectionally.  Primer: Oligo
                     dT. Whole skin from 11 week old C57BL/6 female mice.
                     Average insert size: 1.0 kb; Uni-ZAP XR Vector; ~5'
                     adaptor sequence: 5' GAATTCGGCACGAG 3' ~3' adaptor
                     sequence: 5' CTCGAGTTTTTTTTTTTTTTTTT 3'"
BASE COUNT        3 a     13 c     12 g     12 t
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 9;  Length 40;
  Best Local Similarity   80.0%;  Pred. No. 1.5e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 GCTGAGAAACCTTCTTTTGC 20
          |||| ||| ||| || ||||
Db      9 GCTGTGAATCCTGCTGTTGC 28


RESULT 9
AU102560/c
LOCUS       AU102560                 50 bp    mRNA    linear   EST 30-AUG-2001
DEFINITION  AU102560 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
            CAS01068, mRNA sequence.
ACCESSION   AU102560
VERSION     AU102560.1  GI:13552081
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
```

```
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Suzuki,Y., Taira,H., Tsunoda,T., Mizushima-Sugano,J., Sese,J., Hata
            ,H., Ota,T., Isogai,T., Tanaka,T., Morishita,S., Okubo,K., Sakaki
            ,Y., Nakamura,Y., Suyama,A. and Sugano,S.
  TITLE     Diverse transcriptional initiation revealed by fine, large-scale
            mapping of mRNA start sites
  JOURNAL   EMBO Rep. 2 (5), 388-393 (2001)
  MEDLINE   21270072
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp
            Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano
            ,S. Construction and characterization of a full length-enriched and
            a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).
FEATURES              Location/Qualifiers
     source           1. .50
                      /organism="Homo sapiens"
                      /db_xref="taxon:9606"
                      /clone="CAS01068"
                      /clone_lib="Sugano Homo sapiens cDNA library"
                      /note="Differential display comparison of untreated and
                      dimethylfumarate treated U937 cells"
BASE COUNT       12 a     17 c     16 g      5 t
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 9;  Length 50;
  Best Local Similarity   80.0%;  Pred. No. 1.6e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 GCTGAGAAACCTTCTTTTGC 20
          ||||||  ||| ||| ||||
Db     42 GCTGAGGCACCGTCTATTGC 23


RESULT 10
AU102565/c
LOCUS       AU102565                 50 bp    mRNA    linear   EST 30-AUG-2001
DEFINITION  AU102565 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
            HEP07926, mRNA sequence.
ACCESSION   AU102565
VERSION     AU102565.1  GI:13552086
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Suzuki,Y., Taira,H., Tsunoda,T., Mizushima-Sugano,J., Sese,J., Hata
            ,H., Ota,T., Isogai,T., Tanaka,T., Morishita,S., Okubo,K., Sakaki
            ,Y., Nakamura,Y., Suyama,A. and Sugano,S.
  TITLE     Diverse transcriptional initiation revealed by fine, large-scale
            mapping of mRNA start sites
  JOURNAL   EMBO Rep. 2 (5), 388-393 (2001)
  MEDLINE   21270072
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp
            Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano
            ,S. Construction and characterization of a full length-enriched and
            a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).
FEATURES              Location/Qualifiers
     source           1. .50
                      /organism="Homo sapiens"
                      /db_xref="taxon:9606"
                      /clone="HEP07926"
                      /clone_lib="Sugano Homo sapiens cDNA library"
                      /note="Differential display comparison of untreated and
                      dimethylfumarate treated U937 cells"
BASE COUNT        9 a     14 c     16 g     11 t
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 9;  Length 50;
  Best Local Similarity   80.0%;  Pred. No. 1.6e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 GCTGAGAAACCTTCTTTTGC 20
          ||||||  ||| ||| ||||
Db     21 GCTGAGGCACCGTCTATTGC 2


RESULT 11
B35044/c
LOCUS       B35044                   76 bp    DNA     linear   GSS 17-OCT-1997
DEFINITION  HS-1026-B1-C09-MR.abi CIT Human Genomic Sperm Library C Homo
            sapiens genomic clone Plate=CT 805 Col=17 Row=F, DNA sequence.
ACCESSION   B35044
VERSION     B35044.1  GI:2534413
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 76)
  AUTHORS   Mahairas,G.G., Zackrone,K.D., Smith,T., Tipton,S., Schmidt,S.,
            Traicoff,R., Abajian,C., Blanchard,A., West,A. and Hood,L.E.
  TITLE     Construction of a Characterized Clone Resource for Genomic
            Sequencing: Generation and Preliminary Analysis of 20,000 Sequence
            Tagged Connectors
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Mahairas GG, Zackrone KD, Hood L
            University of Washington
            Seattle, WA 98195, USA
            Tel: (206) 616-8744
            Fax: (206) 685-7301
            Email: kzackron@u.washington.edu
            Sequence Tagged Connector
            Plate: CT 805  row: F  column: 17
            Class: BAC ends
            High quality sequence stop: 76.
FEATURES              Location/Qualifiers
     source           1. .76
                      /organism="Homo sapiens"
                      /db_xref="taxon:9606"
                      /clone="Plate=CT 805 Col=17 Row=F"
                      /clone_lib="CIT Human Genomic Sperm Library C"
                      /sex="M"
                      /note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in
                      E-Coli DH10B"
BASE COUNT       20 a     18 c     16 g     22 t
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 17;  Length 76;
  Best Local Similarity   80.0%;  Pred. No. 1.7e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 GCTGAGAAACCTTCTTTTGC 20
          |||||||||  |||  ||||
Db     67 GCTGAGAAAATTTCCATTGC 48


RESULT 12
BQ568492/c
LOCUS       BQ568492                 66 bp    mRNA    linear   EST 19-JUN-2002
DEFINITION  gi109b06.y1 Mouse Organ of Corti cDNA pBluescript Mus musculus cDNA
            clone gi109b06 5', mRNA sequence.
ACCESSION   BQ568492
VERSION     BQ568492.1  GI:21471809
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 66)
AUTHORS Kachar,B.
TITLE EST analysis of gene expression in the mouse Organ of Corti at the onset of hearing
JOURNAL Unpublished (2002)
COMMENT Contact: Kachar,B.
Structural Cell Biology
National Institute of Deafness and other Communication Disorders
50/4249 South Drive, NIH, Bethesda, MD 20892-8027, USA
Tel: 301-402-1599
Fax: 301-402-1765
Email: kacharb@nidcd.nih.gov
Plate: 109 row: b column: 06
Seq primer: M13RP1 reverse primer (ABI).
FEATURES Location/Qualifiers
source 1. .66
/organism="Mus musculus"
/strain="BALB/c"
/db_xref="taxon:10090"
/clone="gi109b06"
/clone_lib="Mouse Organ of Corti cDNA pBluescript"
/sex="male and female"
/dev_stage="Post natal day 5 to 13"
/note="Organ: Organ of Corti; Vector: pBluescript; The organ of Corti (OC) was fine dissected from a total of 386 OC as follows: 102 samples from post-natal (P) day 5; 72 from P6; 60 from P7; 46 from P8; 18 from P9; 20 from P10; 14 from P12 and 24 from P13. After killing animals by cervical dislocation followed by decapitation, the bulla was removed and opened in Leibowitz medium. The bony capsule of the cochlea was chipped away, stria vascularis and spiral ligament were removed and the sensory epithelium was carefully dissected out of the modiolus. Total RNA was extracted using the micro Fasttrack kit (catalog # K1593-02; Invitrogen, Carlsbad, CA), according to manufacturer's instructions. Reverse transcription and library construction were carried out with the Uni-Zap XR vector kit (catalog # 237211, Stratagene) and Uni-Zap XR Gigapack III Gold Cloning kit (catalog # 237612), both from Stratagene (La Jolla, CA, USA), according to manufacturer's instructions. Briefly: 1.5 ug mRNA was reverse transcribed using a hybrid oligo(dT) linker-primer that contains an Xho I site. First strand synthesis was primed with the linker- primer and transcribed using Moloney murine leukemia virus reverse transcriptase (MMLV-RT) and 5-methyl dCTP. The second strand was synthesized with DNA polymerase and RNase H. Complementary DNA was blunt ended with Pfu DNA polymerase, ligated with EcoR I adapters in the presence of ligase and digested with Xho I. The cDNA was sequentially size fractionated over Pharmacia Size Sep400 (Pharmacia, Uppsala, Sweden) and Clontech Chroma Spin-1000 (Clontech, Palo Alto, CA) columns to enrich for cDNAs greater than 400bp and 1000 bp , respectively. The cDNA was then directionally ligated to the Uni-ZAP XR vector, which had been predigested with EcoR I and Xho I. The phagemid was packaged with Gigapak III Gold and, upon titration on XL1 Blue MRF' cells, the yield of the phage library was estimated to be 11,100,000 recombinants. Stratagene's ExAssist Interference resistance helper phage (catalogue # 211203) was adopted to rescue plasmid DNA from the phages. Upon plating of the rescued library, individual cDNA clones were selected and grown in 96-well, 2 ml growth plate. Plasmid DNA was purified from 200 ul of saturated culture with the Concert96(TM) plasmid purification kit (Invitrogen, Carlsbad, CA) as instructed by the manufacturer. ESTs from the 5' end of the cDNA clones were generated with the universal M13 reverse primer (CAGGAAACAGCTATGACC) and 25% strength BigDye terminator sequencing chemistry (Applied Biosystems, Foster City, CA). Sequencing reactions were performed on MJ Tetrad thermal cyclers (MJ Research, Waltham, MA), and analyzed on 3700 automated capillary sequencers using POP5 polymer (Applied Biosystems, Foster City, CA). The frequency distribution of the library is as follows: 72% of genes have 1 copy; 14.3% 2; 12% 3-10; 1.4% 11-50 and 0.1% 51-150. As to gene function, 45% of genes are present in GenBank and have know function; 23% have hits in GenBank, but do not have assigned function; 12% are uncharacterized ESTs and 20% are unidentified."
BASE COUNT 22 a 14 c 15 g 15 t
ORIGIN

Query Match 67.0%; Score 13.4; DB 14; Length 66;
Best Local Similarity 93.3%; Pred. No. 2.1e+04;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 TGAGAAACCTTCTTT 17
||||||| |||||||
Db 50 TGAGAAAGCTTCTTT 36

RESULT 13
AA512483/c
LOCUS AA512483 76 bp mRNA linear EST 08-JUL-1997
DEFINITION vj16h04.r1 Barstead mouse irradiated colon MPLRB7 Mus musculus cDNA clone IMAGE:921943 5', mRNA sequence.
ACCESSION AA512483
VERSION AA512483.1 GI:2250337
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 76)
AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.
TITLE The WashU-HHMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Marra M/Mouse EST Project
WashU-HHMI Mouse EST Project
Washington University School of MedicineP
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:534159
Seq primer: -28m13 rev2 ET from Amersham
High quality sequence stop: 65.
FEATURES Location/Qualifiers
source 1. .76
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:921943"
/clone_lib="Barstead mouse irradiated colon MPLRB7"
/dev_stage="8 weeks"
/lab_host="DH10B"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: EcoRI; Site_2: NotI; Tissue obtained from 8 week old mouse. Colon was harvested 72 hours after irradiation with 1400 Gys. 1st strand cDNA was primed with a Not I - oligo(dT) primer [5'TGTTACGAATCTGAAGTGGGAGCGGCCGCCCTTTTTTTTTTTTTTTTTTTTTTT T 3']; double-stranded cDNA was ligated to Eco RI adaptors [AATTCGGATCCTTG], digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library constructed by Bob Barstead. "
BASE COUNT 26 a 17 c 22 g 11 t
ORIGIN

```
  Query Match             67.0%;  Score 13.4;  DB 9;  Length 76;
  Best Local Similarity   93.3%;  Pred. No. 2.2e+04;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      6 GAAACCTTCTTTTGC 20
          || ||||||||||||
Db     39 GATACCTTCTTTTGC 25


RESULT 14
CNS03DGG/c
LOCUS       CNS03DGG                  64 bp    DNA     linear   GSS 15-MAY-2000
DEFINITION  Tetraodon nigroviridis genome survey sequence PUC-Ori end of clone
            016L24 of library G from Tetraodon nigroviridis, genomic survey
            sequence.
ACCESSION   AL239065
VERSION     AL239065.1  GI:7898200
KEYWORDS    GSS; genome survey sequence.
SOURCE      Tetraodon nigroviridis.
  ORGANISM  Tetraodon nigroviridis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
            Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
            Tetraodontidae; Tetraodon.
REFERENCE   1  (bases 1 to 64)
  AUTHORS   Roest-Crollius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C.,
            Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F.,
            Saurin,W. and Weissenbach,J.
  TITLE     Human gene number estimate provided by genome wide analysis using
            Tetraodon nigroviridis DNA sequence
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 64)
  AUTHORS   Roest-Crollius,H., Jaillon,O., Dasilva,C., Fizames,C., Fisher,C.,
            Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and
            Weissenbach,J.
  TITLE     Charaterization and repeat analysis of the compact genome of the
            freshwater pufferfish Tetraodon nigroviridis
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 64)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-APR-2000)
COMMENT     This sequence is a single read and was generated as part of a large
            scale clone-end sequencing project of the Tetraodon nigroviridis
            genome. For more information, please take a look at
            http://www.genoscope.cns.fr/Tetraodon.
FEATURES             Location/Qualifiers
     source          1. .64
                     /organism="Tetraodon nigroviridis"
                     /db_xref="taxon:99883"
                     /clone="016L24"
                     /clone_lib="G"
                     /note="Genoscope sequence ID : C0BG016DF12SP1~end :
                     PUC-Ori"
BASE COUNT        9 a      9 c     23 g     21 t      2 others
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 17;  Length 64;
  Best Local Similarity   78.9%;  Pred. No. 2.6e+04;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      2 CTGAGAAACCTTCTTTTGC 20
          || ||||||| |||||  |
Db     19 CTCAGAAACCATCTTTNCC 1


RESULT 15
AA103844/c
LOCUS       AA103844                  72 bp    mRNA    linear   EST 29-OCT-1996
DEFINITION  mo44c05.r1 Life Tech mouse embryo 15 5dpc 10667012 Mus musculus
            cDNA clone IMAGE:556424 5', mRNA sequence.
ACCESSION   AA103844
VERSION     AA103844.1  GI:1650005
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 72)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:337216
            Seq primer: -28M13 rev1 from Amersham
            High quality sequence stop: 55.
FEATURES             Location/Qualifiers
     source          1. .72
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:556424"
                     /clone_lib="Life Tech mouse embryo 15 5dpc 10667012"
                     /tissue_type="embryo"
                     /dev_stage="15.5dpc embryos"
                     /lab_host="DH10B"
                     /note="Organ: whole embryo; Vector: pCMV-SPORT2; Site_1:
                     SalI; Site_2: NotI; Cloned unidirectionally.  Primer:
                     Oligo dT.  15.5dpc embryos. pCMV-SPORT2 vector."
BASE COUNT       18 a     22 c     19 g     13 t
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 9;  Length 72;
  Best Local Similarity   83.3%;  Pred. No. 2.7e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      3 TGAGAAACCTTCTTTTGC 20
          ||||| |||||||| | |
Db     66 TGAGACACCTTCTTCTCC 49


Search completed: November 26, 2002, 17:57:08
Job time : 813.5 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 3, 2002, 18:14:22 ; Search time 351.3 Seconds
(without alignments)
1656.863 Million cell updates/sec

Title: US-09-296-264-22
Perfect score: 20
Sequence: 1 aacatctgtggggttggtgt 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 995600

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
c	1	19.6	98.0	21	6	AX153911	AX153911 Sequence
c	2	15.8	79.0	81	6	A74503	A74503 Sequence 18
c	3	15.8	79.0	81	6	A77482	A77482 Sequence 18
	4	14.8	74.0	39	10	MMTCRVJAA	Z22829 M.musculus
	5	14.2	71.0	22	6	AX203436	AX203436 Sequence
	6	14.2	71.0	55	14	HIV1U45115	U45115 Human immun
	7	14.2	71.0	55	14	HIV1U45141	U45141 Human immun
	8	14.2	71.0	55	14	HIV1U45143	U45143 Human immun
	9	14.2	71.0	55	14	HIV1U45145	U45145 Human immun
	10	14.2	71.0	55	14	HIV1U45159	U45159 Human immun
	11	14.2	71.0	55	14	HIV1U45167	U45167 Human immun
	12	14.2	71.0	55	14	HIV1U45175	U45175 Human immun
	13	13.8	69.0	18	6	AX353251	AX353251 Sequence
	14	13.8	69.0	18	6	AX353260	AX353260 Sequence
	15	13.8	69.0	18	6	AX353281	AX353281 Sequence
	16	13.8	69.0	18	6	AX363096	AX363096 Sequence
	17	13.8	69.0	18	6	AX363105	AX363105 Sequence
	18	13.8	69.0	18	6	AX363126	AX363126 Sequence
c	19	13.8	69.0	32	6	AR208607	AR208607 Sequence
c	20	13.8	69.0	33	6	AX280425	AX280425 Sequence
	21	13.8	69.0	73	6	I92349	I92349 Sequence 1
	22	13.6	68.0	36	6	AR120384	AR120384 Sequence
c	23	13.6	68.0	80	6	A52215	A52215 Sequence 5
	24	13.4	67.0	29	6	I56811	I56811 Sequence 1
	25	13.4	67.0	66	6	AR012567	AR012567 Sequence
	26	13.4	67.0	66	6	AR020395	AR020395 Sequence
	27	13.4	67.0	66	6	AR109416	AR109416 Sequence
	28	13.4	67.0	66	6	I82741	I82741 Sequence 18
	29	13.4	67.0	94	6	AX470106	AX470106 Sequence
	30	13.4	67.0	94	6	AX474165	AX474165 Sequence
	31	13.2	66.0	18	6	A69615	A69615 Sequence 24
	32	13.2	66.0	20	6	AX012555	AX012555 Sequence
	33	13.2	66.0	29	6	AX082659	AX082659 Sequence
c	34	13.2	66.0	30	6	AR193713	AR193713 Sequence
c	35	13.2	66.0	33	10	MMTCRJB11	X80372 M.musculus
c	36	13.2	66.0	44	6	AX297630	AX297630 Sequence
	37	13.2	66.0	69	6	AR097013	AR097013 Sequence
c	38	13.2	66.0	70	6	AR123744	AR123744 Sequence
c	39	13.2	66.0	74	6	AR097014	AR097014 Sequence
	40	13.2	66.0	79	6	AR097015	AR097015 Sequence
	41	13.2	66.0	84	14	POLTER5B	M24195 Poliovirus
c	42	13.2	66.0	85	9	AB033905S04	AB033903 Homo sapi
c	43	13.2	66.0	95	6	AR097016	AR097016 Sequence
	44	13.2	66.0	100	6	I91488	I91488 Sequence 22
	45	12.8	64.0	18	6	AX353290	AX353290 Sequence

ALIGNMENTS

RESULT 1
AX153911/c
LOCUS AX153911 21 bp DNA linear PAT 22-JUN-2001
DEFINITION Sequence 9 from Patent WO0138576.
ACCESSION AX153911
VERSION AX153911.1 GI:14535525
KEYWORDS .
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 21)
AUTHORS Cargill,M., Ireland,J.S. and Lander,E.S.
TITLE Human single nucleotide polymorphisms
JOURNAL Patent: WO 0138576-A 9 31-MAY-2001;

```
          WHITEHEAD INSTITUTE FOR BIOMEDICAL RESEARCH (US)
FEATURES             Location/Qualifiers
     source          1. .21
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT        7 a      7 c      3 g      3 t      1 others
ORIGIN

  Query Match             98.0%;  Score 19.6;  DB 6;  Length 21;
  Best Local Similarity   95.0%;  Pred. No. 5.3;
  Matches   19;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 AACATCTGTGGGGTTGGTGT 20
          |||||||||:||||||||||
Db     20 AACATCTGTRGGGTTGGTGT 1


RESULT 2
A74503/c
LOCUS       A74503                   81 bp    DNA     linear   PAT 15-OCT-1999
DEFINITION  Sequence 189 from Patent WO9401548.
ACCESSION   A74503
VERSION     A74503.1  GI:6064517
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 81)
  AUTHORS   Sibson,D.R. and Gross,J.
  TITLE     HUMAN NUCLEIC ACID FRAGMENTS, ISOLATED FROM BRAIN ADRENAL TISSUE,
            PLACENTA OR BONE NARROW
  JOURNAL   Patent: WO 9401548-A 189 20-JAN-1994;
            MEDICAL RES COUNCIL (GB); SIBSON DAVID ROSS (GB)
FEATURES             Location/Qualifiers
     source          1. .81
                     /organism="unidentified"
                     /db_xref="taxon:32644"
BASE COUNT       28 a     22 c     13 g     18 t
ORIGIN

  Query Match             79.0%;  Score 15.8;  DB 6;  Length 81;
  Best Local Similarity   89.5%;  Pred. No. 7.9e+02;
  Matches   17;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      2 ACATCTGTGGGGTTGGTGT 20
          |||| | ||||||||||||
Db     42 ACATTTTTGGGGTTGGTGT 24


RESULT 3
A77482/c
LOCUS       A77482                   81 bp    DNA     linear   PAT 19-OCT-1999
DEFINITION  Sequence 189 from Patent EP0587279.
ACCESSION   A77482
VERSION     A77482.1  GI:6089147
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 81)
  AUTHORS   Sibson,D.R. and Hadfield,K.M.
  TITLE     HUMAN NUCLEIC ACID FRAGMENTS ISOLATED FROM BRAIN, ADRENAL TISSUE,
            PLACENTA OR BONE NARROW AND THEIR USE
  JOURNAL   Patent: EP 0587279-A 189 16-MAR-1994;
            MEDICAL RES COUNCIL (GB)
FEATURES             Location/Qualifiers
     source          1. .81
                     /organism="unidentified"
                     /db_xref="taxon:32644"
BASE COUNT       28 a     22 c     13 g     18 t
ORIGIN

  Query Match             79.0%;  Score 15.8;  DB 6;  Length 81;
  Best Local Similarity   89.5%;  Pred. No. 7.9e+02;
  Matches   17;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      2 ACATCTGTGGGGTTGGTGT 20
          |||| | ||||||||||||
Db     42 ACATTTTTGGGGTTGGTGT 24


RESULT 4
MMTCRVJAA
LOCUS       MMTCRVJAA                39 bp    mRNA    linear   ROD 01-JAN-1994
DEFINITION  M.musculus T cell receptor V alpha 34S.281, and T cell receptor J
            alpha LB2 mRNA.
ACCESSION   Z22829
VERSION     Z22829.1  GI:438165
KEYWORDS    j_segment; n_region; T cell receptor; T cell receptor J alpha LB2
            gene; T cell receptor V alpha 34S-281 gene; V_region.
SOURCE      Mus musculus.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 39)
  AUTHORS   Roger,T.
  TITLE     Correlation between Vb gene usage and expression of Vg1-Cg4 or
            Vg2-Cg2 mRNA in Vb8.2+ autoreactive ab T cells from normal mice
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 39)
  AUTHORS   Roger,T.T.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-MAY-1993) THIERRY TR ROGER -, Lab.
            d'Immunodifferenciation, UNIVERSITE DENIS, DIDEROT (PARIS 7), HALL
            DES BIOTECHNOLOGIES, 2, place Jussieu, PARIS, PARIS, 75251, PARIS
            Cedex 05, FRANCE
FEATURES             Location/Qualifiers
     source          1. .39
                     /organism="Mus musculus"
                     /strain="DBA/2"
                     /db_xref="taxon:10090"
                     /chromosome="14"
                     /cell_line="1G3"
                     /cell_type="CD4-CD8-T lymphocyte"
                     /tissue_type="spleen"
                     /dev_stage="cell expansion stage"
     gene            1. .10
                     /gene="T cell receptor V alpha 34S.281"
     exon            1. .10
                     /gene="T cell receptor V alpha 34S.281"
     V_region        1. .10
                     /gene="T cell receptor V alpha 34S.281"
     N_region        11. .16
     gene            17. .39
                     /gene="T cell receptor J alpha LB2"
     exon            17. .39
                     /gene="T cell receptor J alpha LB2"
     J_segment       17. .39
                     /gene="T cell receptor J alpha LB2"
BASE COUNT        7 a      5 c     15 g     12 t
ORIGIN

  Query Match             74.0%;  Score 14.8;  DB 10;  Length 39;
  Best Local Similarity   88.9%;  Pred. No. 2.9e+03;
  Matches   16;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      2 ACATCTGTGGGGTTGGTG 19
          ||||||||| ||| ||||
Db      2 ACATCTGTGTGGTGGGTG 19


RESULT 5
AX203436
LOCUS       AX203436                 22 bp    DNA     linear   PAT 30-AUG-2001
DEFINITION  Sequence 66 from Patent WO0153520.
ACCESSION   AX203436
```

```
VERSION     AX203436.1  GI:15392837
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 22)
  AUTHORS   Cullen,P. and Seedorf,U.
  TITLE     Gene chip for neonate screening
  JOURNAL   Patent: WO 0153520-A 66 26-JUL-2001;
            Cullen, Paul (DE) ; Seedorf, Udo (DE)
FEATURES             Location/Qualifiers
     source          1. .22
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT        2 a      4 c     10 g      6 t
ORIGIN

  Query Match            71.0%;  Score 14.2;  DB 6;  Length 22;
  Best Local Similarity  84.2%;  Pred. No. 6.2e+03;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 ACATCTGTGGGGTTGGTGT 20
          |||||| ||||| ||| ||
Db      4 ACATCTTTGGGGCTGGCGT 22


RESULT 6
HIV1U45115
LOCUS       HIV1U45115               55 bp    DNA     linear   VRL 24-JUL-1996
DEFINITION  Human immunodeficiency virus type 1, patient p74, fragment 117.1
            region B, reverse transcriptase (pol) gene, partial cds.
ACCESSION   U45115
VERSION     U45115.1  GI:1304843
KEYWORDS    .
SOURCE      Human immunodeficiency virus type 1.
  ORGANISM  Human immunodeficiency virus type 1
            Viruses; Retroid viruses; Retroviridae; Lentivirus; Primate
            lentivirus group.
REFERENCE   1  (bases 1 to 55)
  AUTHORS   Cleland,A., Watson,H.G., Robertson,P., Ludlam,C.A. and Brown,A.J.
  TITLE     Evolution of zidovudine resistance-associated genotypes in human
            immunodeficiency virus type 1-infected patients
  JOURNAL   J. Acquir. Immune Defic. Syndr. Hum. Retrovirol. 12 (1), 6-18
            (1996)
  MEDLINE   96242958
   PUBMED   8624762
REFERENCE   2  (bases 1 to 55)
  AUTHORS   Leigh Brown,A.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-JAN-1996) Andrew J. Leigh Brown, University of
            Edinburgh, West Mains Road, Centre for HIV Research ICAPB,
            Edinburgh, EH9 3JN, Scotland
FEATURES             Location/Qualifiers
     source          1. .55
                     /organism="Human immunodeficiency virus type 1"
                     /proviral
                     /isolate="patient p74, PCR"
                     /db_xref="taxon:11676"
                     /map="pos. 3167-3222 HIV-LAI genome"
                     /note="sample 09/89, PBMC; fragment 117.1 region B"
     gene            1. .55
                     /gene="pol"
     CDS             <1. .>55
                     /gene="pol"
                     /codon_start=3
                     /product="reverse transcriptase"
                     /protein_id="AAB04266.1"
                     /db_xref="GI:1304844"
                     /translation="HLWRWGFYTPDKKHQKEP"
BASE COUNT       24 a     11 c     11 g      9 t
ORIGIN

  Query Match            71.0%;  Score 14.2;  DB 14;  Length 55;
  Best Local Similarity  84.2%;  Pred. No. 6.3e+03;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 AACATCTGTGGGGTTGGTG 19
          |||||||||| | ||| |
Db      1 AACATCTGTGGAGGTGGGG 19


RESULT 7
HIV1U45141
LOCUS       HIV1U45141               55 bp    DNA     linear   VRL 24-JUL-1996
DEFINITION  Human immunodeficiency virus type 1, patient p74, fragment 133.2
            region B, reverse transcriptase (pol) gene, partial cds.
ACCESSION   U45141
VERSION     U45141.1  GI:1304895
KEYWORDS    .
SOURCE      Human immunodeficiency virus type 1.
  ORGANISM  Human immunodeficiency virus type 1
            Viruses; Retroid viruses; Retroviridae; Lentivirus; Primate
            lentivirus group.
REFERENCE   1  (bases 1 to 55)
  AUTHORS   Cleland,A., Watson,H.G., Robertson,P., Ludlam,C.A. and Brown,A.J.
  TITLE     Evolution of zidovudine resistance-associated genotypes in human
            immunodeficiency virus type 1-infected patients
  JOURNAL   J. Acquir. Immune Defic. Syndr. Hum. Retrovirol. 12 (1), 6-18
            (1996)
  MEDLINE   96242958
   PUBMED   8624762
REFERENCE   2  (bases 1 to 55)
  AUTHORS   Leigh Brown,A.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-JAN-1996) Andrew J. Leigh Brown, University of
            Edinburgh, West Mains Road, Centre for HIV Research ICAPB,
            Edinburgh, EH9 3JN, Scotland
FEATURES             Location/Qualifiers
     source          1. .55
                     /organism="Human immunodeficiency virus type 1"
                     /proviral
                     /isolate="patient p74, PCR"
                     /db_xref="taxon:11676"
                     /map="pos. 3167-3222 HIV-LAI genome"
                     /note="sample 03/90, PBMC; fragment 133.2 region B"
     gene            1. .55
                     /gene="pol"
     CDS             <1. .>55
                     /gene="pol"
                     /codon_start=3
                     /product="reverse transcriptase"
                     /protein_id="AAB04292.1"
                     /db_xref="GI:1304896"
                     /translation="HLWRWGFYTPDKKHQKEP"
BASE COUNT       24 a     11 c     11 g      9 t
ORIGIN

  Query Match            71.0%;  Score 14.2;  DB 14;  Length 55;
  Best Local Similarity  84.2%;  Pred. No. 6.3e+03;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 AACATCTGTGGGGTTGGTG 19
          |||||||||| | ||| |
Db      1 AACATCTGTGGAGGTGGGG 19


RESULT 8
HIV1U45143
LOCUS       HIV1U45143               55 bp    DNA     linear   VRL 24-JUL-1996
DEFINITION  Human immunodeficiency virus type 1, patient p74, fragment 133.3
            region B, reverse transcriptase (pol) gene, partial cds.
ACCESSION   U45143
VERSION     U45143.1  GI:1304899
KEYWORDS    .
SOURCE      Human immunodeficiency virus type 1.
```

```
  ORGANISM  Human immunodeficiency virus type 1
            Viruses; Retroid viruses; Retroviridae; Lentivirus; Primate
            lentivirus group.
REFERENCE   1  (bases 1 to 55)
  AUTHORS   Cleland,A., Watson,H.G., Robertson,P., Ludlam,C.A. and Brown,A.J.
  TITLE     Evolution of zidovudine resistance-associated genotypes in human
            immunodeficiency virus type 1-infected patients
  JOURNAL   J. Acquir. Immune Defic. Syndr. Hum. Retrovirol. 12 (1), 6-18
            (1996)
  MEDLINE   96242958
   PUBMED   8624762
REFERENCE   2  (bases 1 to 55)
  AUTHORS   Leigh Brown,A.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-JAN-1996) Andrew J. Leigh Brown, University of
            Edinburgh, West Mains Road, Centre for HIV Research ICAPB,
            Edinburgh, EH9 3JN, Scotland
FEATURES             Location/Qualifiers
     source          1. .55
                     /organism="Human immunodeficiency virus type 1"
                     /proviral
                     /isolate="patient p74, PCR"
                     /db_xref="taxon:11676"
                     /map="pos. 3167-3222 HIV-LAI genome"
                     /note="sample 03/90, PBMC; fragment 133.3 region B"
     gene            1. .55
                     /gene="pol"
     CDS             <1. .>55
                     /gene="pol"
                     /codon_start=3
                     /product="reverse transcriptase"
                     /protein_id="AAB04294.1"
                     /db_xref="GI:1304900"
                     /translation="HLWRWGFYTPDKKHQKEP"
BASE COUNT       24 a     11 c     11 g      9 t
ORIGIN

  Query Match             71.0%;  Score 14.2;  DB 14;  Length 55;
  Best Local Similarity   84.2%;  Pred. No. 6.3e+03;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 AACATCTGTGGGGTTGGTG 19
          ||||||||||| | ||| |
Db      1 AACATCTGTGGAGGTGGGG 19


RESULT 9
HIV1U45145
LOCUS       HIV1U45145               55 bp    DNA     linear   VRL 24-JUL-1996
DEFINITION  Human immunodeficiency virus type 1, patient p74, fragment 133.4
            region B, reverse transcriptase (pol) gene, partial cds.
ACCESSION   U45145
VERSION     U45145.1  GI:1304903
KEYWORDS    .
SOURCE      Human immunodeficiency virus type 1.
  ORGANISM  Human immunodeficiency virus type 1
            Viruses; Retroid viruses; Retroviridae; Lentivirus; Primate
            lentivirus group.
REFERENCE   1  (bases 1 to 55)
  AUTHORS   Cleland,A., Watson,H.G., Robertson,P., Ludlam,C.A. and Brown,A.J.
  TITLE     Evolution of zidovudine resistance-associated genotypes in human
            immunodeficiency virus type 1-infected patients
  JOURNAL   J. Acquir. Immune Defic. Syndr. Hum. Retrovirol. 12 (1), 6-18
            (1996)
  MEDLINE   96242958
   PUBMED   8624762
REFERENCE   2  (bases 1 to 55)
  AUTHORS   Leigh Brown,A.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-JAN-1996) Andrew J. Leigh Brown, University of
            Edinburgh, West Mains Road, Centre for HIV Research ICAPB,
            Edinburgh, EH9 3JN, Scotland
FEATURES             Location/Qualifiers
     source          1. .55
                     /organism="Human immunodeficiency virus type 1"
                     /proviral
                     /isolate="patient p74, PCR"
                     /db_xref="taxon:11676"
                     /map="pos. 3167-3222 HIV-LAI genome"
                     /note="sample 03/90, PBMC; fragment 133.4 region B"
     gene            1. .55
                     /gene="pol"
     CDS             <1. .>55
                     /gene="pol"
                     /codon_start=3
                     /product="reverse transcriptase"
                     /protein_id="AAB04296.1"
                     /db_xref="GI:1304904"
                     /translation="HLWRWGFYTPDKKHQKEP"
BASE COUNT       24 a     11 c     11 g      9 t
ORIGIN

  Query Match             71.0%;  Score 14.2;  DB 14;  Length 55;
  Best Local Similarity   84.2%;  Pred. No. 6.3e+03;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 AACATCTGTGGGGTTGGTG 19
          ||||||||||| | ||| |
Db      1 AACATCTGTGGAGGTGGGG 19


RESULT 10
HIV1U45159
LOCUS       HIV1U45159               55 bp    DNA     linear   VRL 24-JUL-1996
DEFINITION  Human immunodeficiency virus type 1, patient p74, fragment 133.11
            region B, reverse transcriptase (pol) gene, partial cds.
ACCESSION   U45159
VERSION     U45159.1  GI:1304931
KEYWORDS    .
SOURCE      Human immunodeficiency virus type 1.
  ORGANISM  Human immunodeficiency virus type 1
            Viruses; Retroid viruses; Retroviridae; Lentivirus; Primate
            lentivirus group.
REFERENCE   1  (bases 1 to 55)
  AUTHORS   Cleland,A., Watson,H.G., Robertson,P., Ludlam,C.A. and Brown,A.J.
  TITLE     Evolution of zidovudine resistance-associated genotypes in human
            immunodeficiency virus type 1-infected patients
  JOURNAL   J. Acquir. Immune Defic. Syndr. Hum. Retrovirol. 12 (1), 6-18
            (1996)
  MEDLINE   96242958
   PUBMED   8624762
REFERENCE   2  (bases 1 to 55)
  AUTHORS   Leigh Brown,A.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-JAN-1996) Andrew J. Leigh Brown, University of
            Edinburgh, West Mains Road, Centre for HIV Research ICAPB,
            Edinburgh, EH9 3JN, Scotland
FEATURES             Location/Qualifiers
     source          1. .55
                     /organism="Human immunodeficiency virus type 1"
                     /proviral
                     /isolate="patient p74, PCR"
                     /db_xref="taxon:11676"
                     /map="pos. 3167-3222 HIV-LAI genome"
                     /note="sample date 03/90, PBMC; fragment 133.11 region B"
     gene            1. .55
                     /gene="pol"
     CDS             <1. .>55
                     /gene="pol"
                     /codon_start=3
                     /product="reverse transcriptase"
                     /protein_id="AAB04310.1"
                     /db_xref="GI:1304932"
                     /translation="HLWRWGFYTPDKKHQKEP"
BASE COUNT       24 a     11 c     11 g      9 t
ORIGIN
```

```
  Query Match             71.0%;  Score 14.2;  DB 14;  Length 55;
  Best Local Similarity   84.2%;  Pred. No. 6.3e+03;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 AACATCTGTGGGGTTGGTG 19
          ||||||||||| | ||| |
Db      1 AACATCTGTGGAGGTGGGG 19


RESULT 11
HIV1U45167
LOCUS       HIV1U45167               55 bp    RNA     linear   VRL 24-JUL-1996
DEFINITION  Human immunodeficiency virus type 1, patient p74, fragment c133.4
            region B, reverse transcriptase (pol) gene, partial cds.
ACCESSION   U45167
VERSION     U45167.1  GI:1304947
KEYWORDS    .
SOURCE      Human immunodeficiency virus type 1.
  ORGANISM  Human immunodeficiency virus type 1
            Viruses; Retroid viruses; Retroviridae; Lentivirus; Primate
            lentivirus group.
REFERENCE   1  (bases 1 to 55)
  AUTHORS   Cleland,A., Watson,H.G., Robertson,P., Ludlam,C.A. and Brown,A.J.
  TITLE     Evolution of zidovudine resistance-associated genotypes in human
            immunodeficiency virus type 1-infected patients
  JOURNAL   J. Acquir. Immune Defic. Syndr. Hum. Retrovirol. 12 (1), 6-18
            (1996)
  MEDLINE   96242958
   PUBMED   8624762
REFERENCE   2  (bases 1 to 55)
  AUTHORS   Leigh Brown,A.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-JAN-1996) Andrew J. Leigh Brown, University of
            Edinburgh, West Mains Road, Centre for HIV Research ICAPB,
            Edinburgh, EH9 3JN, Scotland
FEATURES             Location/Qualifiers
     source          1. .55
                     /organism="Human immunodeficiency virus type 1"
                     /isolate="patient p74, PCR"
                     /db_xref="taxon:11676"
                     /map="pos. 3167-3222 HIV-LAI genome"
                     /note="sample date 03/90, plasma; fragment c133.4 region
                     B"
     gene            1. .55
                     /gene="pol"
     CDS             <1. .>55
                     /gene="pol"
                     /codon_start=3
                     /product="reverse transcriptase"
                     /protein_id="AAB04318.1"
                     /db_xref="GI:1304948"
                     /translation="HLWRWGFYTPDKKHQKEP"
BASE COUNT       24 a     11 c     11 g      9 t
ORIGIN

  Query Match             71.0%;  Score 14.2;  DB 14;  Length 55;
  Best Local Similarity   84.2%;  Pred. No. 6.3e+03;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 AACATCTGTGGGGTTGGTG 19
          ||||||||||| | ||| |
Db      1 AACATCTGTGGAGGTGGGG 19


RESULT 12
HIV1U45175
LOCUS       HIV1U45175               55 bp    RNA     linear   VRL 24-JUL-1996
DEFINITION  Human immunodeficiency virus type 1, patient p74, fragment c133.8
            region B, reverse transcriptase (pol) gene, partial cds.
ACCESSION   U45175
VERSION     U45175.1  GI:1304963
KEYWORDS    .
SOURCE      Human immunodeficiency virus type 1.
  ORGANISM  Human immunodeficiency virus type 1
            Viruses; Retroid viruses; Retroviridae; Lentivirus; Primate
            lentivirus group.
REFERENCE   1  (bases 1 to 55)
  AUTHORS   Cleland,A., Watson,H.G., Robertson,P., Ludlam,C.A. and Brown,A.J.
  TITLE     Evolution of zidovudine resistance-associated genotypes in human
            immunodeficiency virus type 1-infected patients
  JOURNAL   J. Acquir. Immune Defic. Syndr. Hum. Retrovirol. 12 (1), 6-18
            (1996)
  MEDLINE   96242958
   PUBMED   8624762
REFERENCE   2  (bases 1 to 55)
  AUTHORS   Leigh Brown,A.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (11-JAN-1996) Andrew J. Leigh Brown, University of
            Edinburgh, West Mains Road, Centre for HIV Research ICAPB,
            Edinburgh, EH9 3JN, Scotland
FEATURES             Location/Qualifiers
     source          1. .55
                     /organism="Human immunodeficiency virus type 1"
                     /isolate="patient p74, PCR"
                     /db_xref="taxon:11676"
                     /map="pos. 3167-3222 HIV-LAI genome"
                     /note="sample date 03/90, plasma; fragment c133.8 region
                     B"
     gene            1. .55
                     /gene="pol"
     CDS             <1. .>55
                     /gene="pol"
                     /codon_start=3
                     /product="reverse transcriptase"
                     /protein_id="AAB04326.1"
                     /db_xref="GI:1304964"
                     /translation="HLWRWGFYTPDKKHQKEP"
BASE COUNT       24 a     11 c     11 g      9 t
ORIGIN

  Query Match             71.0%;  Score 14.2;  DB 14;  Length 55;
  Best Local Similarity   84.2%;  Pred. No. 6.3e+03;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 AACATCTGTGGGGTTGGTG 19
          ||||||||||| | ||| |
Db      1 AACATCTGTGGAGGTGGGG 19


RESULT 13
AX353251
LOCUS       AX353251                 18 bp    DNA     linear   PAT 06-FEB-2002
DEFINITION  Sequence 457 from Patent EP1174518.
ACCESSION   AX353251
VERSION     AX353251.1  GI:18618333
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Loukachov,V.V., van Gemen,B. and Goudsmit,J.
  TITLE     Collection of binding molecules
  JOURNAL   Patent: EP 1174518-A 457 23-JAN-2002;
            Amsterdam Support Diagnostics B.V. (NL)
FEATURES             Location/Qualifiers
     source          1. .18
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="position 210"
BASE COUNT        4 a      2 c      8 g      4 t
ORIGIN

  Query Match             69.0%;  Score 13.8;  DB 6;  Length 18;
  Best Local Similarity   88.2%;  Pred. No. 1e+04;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;
```

```
Qy      1 AACATCTGTGGGGTTGG 17
          ||||||||||| | |||
Db      1 AACATCTGTGGAGGTGG 17


RESULT 14
AX353260
LOCUS       AX353260                 18 bp    DNA     linear   PAT 06-FEB-2002
DEFINITION  Sequence 466 from Patent EP1174518.
ACCESSION   AX353260
VERSION     AX353260.1  GI:18618342
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Loukachov,V.V., van Gemen,B. and Goudsmit,J.
  TITLE     Collection of binding molecules
  JOURNAL   Patent: EP 1174518-A 466 23-JAN-2002;
            Amsterdam Support Diagnostics B.V. (NL)
FEATURES             Location/Qualifiers
     source          1. .18
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="position 210"
BASE COUNT        3 a      2 c      8 g      5 t
ORIGIN

  Query Match             69.0%;  Score 13.8;  DB 6;  Length 18;
  Best Local Similarity   88.2%;  Pred. No. 1e+04;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 AACATCTGTGGGGTTGG 17
          ||||||||| ||| |||
Db      1 AACATCTGTTGGGGTGG 17


RESULT 15
AX353281
LOCUS       AX353281                 18 bp    DNA     linear   PAT 06-FEB-2002
DEFINITION  Sequence 487 from Patent EP1174518.
ACCESSION   AX353281
VERSION     AX353281.1  GI:18618363
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Loukachov,V.V., van Gemen,B. and Goudsmit,J.
  TITLE     Collection of binding molecules
  JOURNAL   Patent: EP 1174518-A 487 23-JAN-2002;
            Amsterdam Support Diagnostics B.V. (NL)
FEATURES             Location/Qualifiers
     source          1. .18
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="position 210"
BASE COUNT        4 a      2 c      6 g      6 t
ORIGIN

  Query Match             69.0%;  Score 13.8;  DB 6;  Length 18;
  Best Local Similarity   88.2%;  Pred. No. 1e+04;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 AACATCTGTGGGGTTGG 17
          ||||||||| | |||||
Db      1 AACATCTGTTGAGTTGG 17


Search completed: December  3, 2002, 22:23:07
Job time : 357.3 secs
```

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 06:29:45 ; Search time 94.1 Seconds
(without alignments)
478.639 Million cell updates/sec

Title: US-09-296-264-22
Perfect score: 20
Sequence: 1 aacatctgtggggttggtgt 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 2390332

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	20	100.0	20	21	AAZ31452	Human neuropilin m
c	2	20	100.0	21	22	AAH62108	Neuropilin 1 (NRP1
c	3	15.8	79.0	81	15	AAQ76589	Human genome fragm
	4	14.2	71.0	22	22	AAH48939	Human CYP17-associ
	5	14	70.0	60	24	ABN46985	Human spliced tran
	6	13.8	69.0	18	24	ABL89235	HIV-1 related bind
	7	13.8	69.0	18	24	ABL89244	HIV-1 related bind
	8	13.8	69.0	18	24	ABL89265	HIV-1 related bind
	9	13.8	69.0	20	22	AAS05663	Rat NCAM reverse P
c	10	13.8	69.0	23	24	ABL50359	Human cancer cell
c	11	13.8	69.0	32	20	AAZ24764	Human CCR5 recepto
c	12	13.8	69.0	33	23	ABI97596	Endogenous human G
	13	13.8	69.0	50	22	AAL29283	Human SNP oligonuc
	14	13.8	69.0	51	22	AAL29893	Human SNP oligonuc
	15	13.8	69.0	73	19	AAV23400	Template extension
	16	13.6	68.0	36	19	AAV27470	Streptococcus pneu
	17	13.6	68.0	36	24	ABQ84938	Streptococcus pneu
c	18	13.6	68.0	50	22	AAL31484	Human SNP oligonuc
c	19	13.6	68.0	60	24	ABN50089	Human spliced tran
c	20	13.6	68.0	80	17	AAT18556	Human chondrocyte
	21	13.4	67.0	29	17	AAT07256	Oligonucleotide in
	22	13.4	67.0	29	17	AAT08788	Cytokine productio
	23	13.4	67.0	40	24	AAL48379	Cytohesin binding
c	24	13.4	67.0	65	24	ABN55801	Mouse spliced tran
	25	13.4	67.0	65	24	ABN56895	Mouse spliced tran
	26	13.4	67.0	66	17	AAT71511	Glioblastoma U251
	27	13.4	67.0	94	24	AAL48337	Cytohesin binding
	28	13.2	66.0	18	19	AAV20968	Human PRCC-TFE3 co
	29	13.2	66.0	20	20	AAZ39257	Probe for typing H
	30	13.2	66.0	20	20	AAX94291	PCR primer used to
	31	13.2	66.0	20	20	AAX93320	PCR primer used to
c	32	13.2	66.0	24	20	AAX02691	E. przewalskii Fox
c	33	13.2	66.0	30	14	AAQ43254	Sequence of 5' RAC
c	34	13.2	66.0	30	15	AAQ71453	AtT 20 murine proh
c	35	13.2	66.0	44	24	ABI82214	p53 mutation detec
	36	13.2	66.0	51	21	AAA77226	Human clone cg4397
	37	13.2	66.0	51	21	AAA77227	Human clone cg4397
	38	13.2	66.0	60	24	ABN43593	Human spliced tran
c	39	13.2	66.0	70	22	AAF29225	Ligand to CD40 lig
	40	13.2	66.0	90	22	ABA72399	Human foetal liver
	41	13.2	66.0	90	22	AAK20823	Human brain expres
	42	13.2	66.0	90	22	AAK46974	Human bone marrow
	43	13.2	66.0	90	22	AAI52811	Probe #21497 used
	44	13.2	66.0	91	20	AAX88430	Poliovirus CAT rep
	45	13	65.0	89	20	AAX54641	Human IL-5 antisen

ALIGNMENTS

RESULT 1
AAZ31452
ID AAZ31452 standard; DNA; 20 BP.
XX
AC AAZ31452;
XX
DT 07-FEB-2000 (first entry)
XX
DE Human neuropilin mRNA specific antisense oligo GTI3623.
XX
KW Neuropilin; human; growth; metastasis; tumor; neovascularisation;
KW cancer; papilloma; diabetic retinopathy; antisense; ss.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9955855-A2.
XX
PD 04-NOV-1999.
XX
PF 23-APR-1999; 99WO-CA00324.
XX
PR 23-APR-1998; 98US-0082791.
XX
PA (GENE-) GENESENSE TECHNOLOGIES INC.
XX
PI Wright JA, Young AH, Lee YS;
XX
DR WPI; 2000-023357/02.
XX
PT Antisense oligonucleotides that inhibit neuropilin expression, useful
PT for treating cancer -

XX
PS Claim 4; Page 17; 57pp; English.
XX
CC Sequences AAZ31431-460 represent antisense oligonucleotides which
CC inhibit human neuropilin expression. The antisense oligonucleotides can
CC be used to inhibit the growth or metastasis of a mammalian tumor and
CC inhibit neovascularisation. The oligonucleotides may be used to treat
CC various forms of cancers or tumors, such as sarcomas, melanomas,
CC adenomas, carcinomas of solid tissue, hypoxic tumors, squamous cell
CC carcinomas of the mouth, throat, larynx and lung, genitourinary cancers
CC such as cervical and bladder cancer, hematopoietic cancers, colon cancer,
CC breast cancer, pancreatic cancer, renal cancer, brain cancer, skin
CC cancer, liver cancer, head and neck cancers, and nervous system cancers,
CC as well as benign lesions such as papillomas. The methods may be used to
CC treat neovascularisation disorders such as diabetic retinopathy, and
CC retinopathy of prematurity and age related macular degeneration.
XX
SQ Sequence 20 BP; 3 A; 2 C; 8 G; 7 T; 0 other;

Query Match 100.0%; Score 20; DB 21; Length 20;
Best Local Similarity 100.0%; Pred. No. 2.5;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AACATCTGTGGGGTTGGTGT 20
||||||||||||||||||||
Db 1 AACATCTGTGGGGTTGGTGT 20

RESULT 2
AAH62108/c
ID AAH62108 standard; DNA; 21 BP.
XX
AC AAH62108;
XX
DT 12-SEP-2001 (first entry)
XX
DE Neuropilin 1 (NRP1) polymorphism containing DNA fragment #9.
XX
KW Single nucleotide polymorphism; SNP; human; cancer; inflammation;
KW heart disease; paternity testing; forensic science; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Variation replace(11,T)
FT /*tag= a
FT /standard_name= "single nucleotide polymorphism"
XX
PN WO200138576-A2.
XX
PD 31-MAY-2001.
XX
PF 17-NOV-2000; 2000WO-US31639.
XX
PR 24-NOV-1999; 99US-0167334.
XX
PA (WHED) WHITEHEAD INST BIOMEDICAL RES.
XX
PI Cargill M, Ireland JS, Lander ES;
XX
DR WPI; 2001-367705/38.
XX
PT New nucleic acid segments of the human genome, particularly from genes
PT including polymorphic sites,for phenotype correlation, forensics,
PT paternity testing, medicine and genetic analysis
XX
PS Claim 1; Page 29; 80pp; English.
XX
CC DNA sequences AAH62100 - AAH62688 represent segments of human genes which
CC contain single nucleotide polymorphisms (SNPs). A method is included in
CC the invention for analysing a nucleic acid sample, which consists of
CC determining the base occupying any one of the polymorphic sites given in
CC the SNP containing sequences. The nucleotide sequences can be used in the
CC diagnosis or monitoring of diseases, such as cancer, inflammation, heart
CC diseases, diseases of the cardiovascular system, and infection by
CC microorganisms. The oligonucleotides are also useful in the manufacture
CC of a medicament for the treatment or prophylaxis of the diseases, and as
CC a pharmaceutical. SNP containing oligonucleotides are useful in
CC applications such as phenotype correlation, forensics, paternity testing,
CC medicine and genetic analysis.
XX
SQ Sequence 21 BP; 7 A; 8 C; 3 G; 3 T; 0 other;

Query Match 100.0%; Score 20; DB 22; Length 21;
Best Local Similarity 100.0%; Pred. No. 2.5;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AACATCTGTGGGGTTGGTGT 20
||||||||||||||||||||
Db 20 AACATCTGTGGGGTTGGTGT 1

RESULT 3
AAQ76589/c
ID AAQ76589 standard; DNA; 81 BP.
XX
AC AAQ76589;
XX
DT 23-SEP-1994 (first entry)
XX
DE Human genome fragment.
XX
KW Brain; placenta; bone marrow; genetic analysis; gene mapping;
KW detection; homology; human; adrenal tissue; ds.
XX
OS Homo sapiens.
XX
PN WO9401548-A.
XX
PD 20-JAN-1994.
XX
PF 13-JUL-1993; 93WO-GB01467.
XX
PR 13-JUL-1992; 92GB-0014857.
XX
PA (MEDI-) MEDICAL RES COUNCIL.
XX
PI Gross J, Hadfield KM, Howells D, Kelly M, Shaw D;
PI Sibson DR, Starkey M;
XX
DR WPI; 1994-035056/04.
XX
PT New nucleic acid fragment encoding gene products - can be used
PT for genetic analysis and mapping
XX
PS Claim 1; Page 110; 616pp; English.
XX
CC Human nucleic acid fragments, isolated from brain adrenal tissue,
CC the placenta or bone marrow comprise any of: (A) a sequence
CC selected from (AAQ76401-Q77613), (B) an allelic variation of a
CC sequence as described in (A), or (C) a sequence complementary
CC to (A) or (B).
CC Preferred sequences exhibit no more than 90% homology to a human
CC sequence known per se.
XX
SQ Sequence 81 BP; 28 A; 22 C; 13 G; 18 T; 0 other;

Query Match 79.0%; Score 15.8; DB 15; Length 81;
Best Local Similarity 89.5%; Pred. No. 3e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 ACATCTGTGGGGTTGGTGT 20
|||| | ||||||||||||
Db 42 ACATTTTTGGGGTTGGTGT 24

```
RESULT 4
AAH48939
ID   AAH48939 standard; DNA; 22 BP.
XX
AC   AAH48939;
XX
DT   12-NOV-2001  (first entry)
XX
DE   Human CYP17-associated primer #4.
XX
KW   Neonate screening; prenatal screening; gene chip; diagnosis;
KW   phenylketonuria; maple syrup disease; galactosemia; homocysteinuria;
KW   medium-chain acyl-CoA-dehydrogenase deficiency; biotinidase deficiency;
KW   familial hypercholesterolemia; familial defective apolipoprotein-B;
KW   cystic fibrosis; Marfan syndrome; Smith-Lemil-Opitz syndrome;
KW   androgenital syndrome; ss.
XX
OS   Homo sapiens.
XX
PN   WO200153520-A2.
XX
PD   26-JUL-2001.
XX
PF   09-JAN-2001; 2001WO-EP00139.
XX
PR   21-JAN-2000; 2000DE-1002446.
XX
PA   (CULL/) CULLEN P.
PA   (SEED/) SEEDORF U.
XX
PI   Cullen P,  Seedorf U;
XX
DR   WPI; 2001-457616/49.
XX
PT   DNA chip, useful for neonatal or prenatal screening for many genetic
PT   diseases simultaneously, carries oligonucleotides complementary to
PT   phenotypically relevant reference sequences -
XX
PS   Claim 4; Page 36; 101pp; German.
XX
CC   This invention describes a novel nucleotide support (A; gene chip) which
CC   carries a selection of oligonucleotides (I) that are identical, or
CC   complementary, to segments of reference sequences relevant to at least
CC   two genetically determined phenotypes. (A) are used for simultaneous
CC   diagnosis of at least two of the following diseases: phenylketonuria
CC   (maple syrup disease), galactosemia, homocysteinuria, biotinidase
CC   deficiency, medium-chain acyl-CoA-dehydrogenase deficiency, familial
CC   hypercholesterolemia, familial defective apolipoprotein-B, cystic
CC   fibrosis, Marfan syndrome, Smith-Lemil-Opitz syndrome and androgenital
CC   syndrome. Specifically they are used in neonatal or prenatal diagnosis.
CC   (A) require a relatively small number of separate hybridization regions
CC   (about 500 for testing for 21 specified disorders), so can be used for
CC   simultaneous testing for many diseases. Testing is quick, inexpensive,
CC   reliable and more sensitive than current physiological methods.
CC   AAH48868-AAH489166 represent oligonucleotides used to illustrate the
CC   method of the invention.
XX
SQ   Sequence 22 BP; 2 A; 4 C; 10 G; 6 T; 0 other;

  Query Match             71.0%;  Score 14.2;  DB 22;  Length 22;
  Best Local Similarity   84.2%;  Pred. No. 1.5e+03;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 ACATCTGTGGGGTTGGTGT 20
           |||||| ||||| ||| ||
Db       4 ACATCTTTGGGGCTGGCGT 22


RESULT 5
ABN46985
ID   ABN46985 standard; DNA; 60 BP.
XX
AC   ABN46985;
XX
DT   15-JUL-2002  (first entry)
XX
DE   Human spliced transcript detection oligonucleotide SEQ ID NO:19733.
XX
KW   Human; mouse; rat; splice transcript; detection; RNA transcript;
KW   splice variant; transcriptome; oligonucleotide library; ss.
XX
OS   Homo sapiens.
XX
PN   WO200210449-A2.
XX
PD   07-FEB-2002.
XX
PF   20-JUL-2001; 2001WO-IB01903.
XX
PR   28-JUL-2000; 2000US-221607P.
PR   02-MAY-2001; 2001US-287724P.
XX
PA   (COMP-) COMPUGEN INC.
XX
PI   Shoshan A,  Wasserman A,  Mintz E,  Mintz L,  Faigler S;
XX
DR   WPI; 2002-257383/30.
XX
PT   New oligonucleotide libraries comprising oligonucleotides which
PT   selectively hybridize to mRNAs transcribed from a transcription unit of
PT   a genome, useful for detecting tissue-, pathology-, and
PT   developmental-specific genes -
XX
PS   Example 1; SEQ ID 19733; 47pp; English.
XX
CC   The present invention describes oligonucleotide libraries for detecting
CC   messenger RNAs that populate a (sub-)transcriptome, where the
CC   (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC   transcription units that populate a genome. The library comprises
CC   several oligonucleotides, each capable of hybridising selectively to a
CC   set of messenger RNAs transcribed from a given transcription unit of
CC   the genome, which encodes one or more messenger RNA splice variants.
CC   The oligonucleotide libraries are useful for detecting mRNAs from a
CC   biological sample, in expression profiling studies, in qualitatively or
CC   quantitatively characterising the corresponding transcriptome, and in
CC   detecting RNA transcripts and splice variants of human or animal
CC   transcriptomes. The libraries may also be used as specialised mini
CC   libraries to detect transcripts of a sub-transcriptome under a
CC   particular biological or pathological state, and so allowing the
CC   detection of tissue- and pathology-specific genes such as those genes
CC   only expressed in specific tissue under a specific pathological
CC   condition; to detect developmental specific genes; and to detect RNA
CC   transcripts and splice variants of a transcriptome of a patient suffering
CC   from a particular disorder. ABN27253 to ABN59589 represent
CC   oligonucleotide sequences from rats, humans and mice, which are used in
CC   the exemplification of the present invention.
CC   N.B. The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 60 BP; 17 A; 12 C; 16 G; 15 T; 0 other;

  Query Match             70.0%;  Score 14;  DB 24;  Length 60;
  Best Local Similarity   100.0%;  Pred. No. 2.1e+03;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 ACATCTGTGGGGTT 15
           ||||||||||||||
Db      27 ACATCTGTGGGGTT 40


RESULT 6
ABL89235
ID   ABL89235 standard; DNA; 18 BP.
XX
AC   ABL89235;
```

XX
DT 22-MAY-2002 (first entry)
XX
DE HIV-1 related binding molecule oligonucleotide sequence SEQ ID NO:457.
XX
KW Binding molecule; HIV-1; human immunodeficiency virus type 1;
KW reverse transcriptase; binding group; ss.
XX
OS Human immunodeficiency virus type 1.
OS Synthetic.
XX
PN EP1174518-A1.
XX
PD 23-JAN-2002.
XX
PF 20-JUL-2000; 2000EP-0202611.
XX
PR 20-JUL-2000; 2000EP-0202611.
XX
PA (AMST-) AMSTERDAM SUPPORT DIAGNOSTICS BV.
XX
PI Loukachov VV, Van Gemen B, Goudsmit J;
XX
DR WPI; 2002-156696/21.
XX
PT Collection of binding groups for determining or typing samples,
PT especially clinical samples, has groups capable to identify essentially
PT all members of the family of nucleic acids of relatively high
PT significance -
XX
PS Disclosure; Page 117; 166pp; English.
XX
CC The present invention describes a collection of binding groups for a
CC family of nucleic acids comprising members of relative high and relative
CC low significance, where the binding groups are selected to be capable to
CC identify, alone or in combination, essentially all members of the family
CC of nucleic acids of relatively high significance. The collection of
CC binding groups is useful for typing of nucleic acid in a clinical sample,
CC by contacting the nucleic acid with the collection and determining
CC whether one or more binding groups bound to the nucleic acid of the
CC sample. This method is useful for determining whether the sample
CC comprises at least a part of a member of relatively high significance of
CC a family of nucleic acids. The collection of binding groups is useful for
CC diagnosing the severity of a disease caused by a pathogen containing a
CC member of a family of nucleic acids. ABL88779 to ABL89321 represent
CC oligonucleotide sequences used in the exemplification of the present
CC invention.
XX
SQ Sequence 18 BP; 4 A; 2 C; 8 G; 4 T; 0 other;

Query Match 69.0%; Score 13.8; DB 24; Length 18;
Best Local Similarity 88.2%; Pred. No. 2.3e+03;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AACATCTGTGGGGTTGG 17
||||||||||| | |||
Db 1 AACATCTGTGGAGGTGG 17

RESULT 7
ABL89244
ID ABL89244 standard; DNA; 18 BP.
XX
AC ABL89244;
XX
DT 22-MAY-2002 (first entry)
XX
DE HIV-1 related binding molecule oligonucleotide sequence SEQ ID NO:466.
XX
KW Binding molecule; HIV-1; human immunodeficiency virus type 1;
KW reverse transcriptase; binding group; ss.
XX
OS Human immunodeficiency virus type 1.
OS Synthetic.
XX
PN EP1174518-A1.
XX
PD 23-JAN-2002.
XX
PF 20-JUL-2000; 2000EP-0202611.
XX
PR 20-JUL-2000; 2000EP-0202611.
XX
PA (AMST-) AMSTERDAM SUPPORT DIAGNOSTICS BV.
XX
PI Loukachov VV, Van Gemen B, Goudsmit J;
XX
DR WPI; 2002-156696/21.
XX
PT Collection of binding groups for determining or typing samples,
PT especially clinical samples, has groups capable to identify essentially
PT all members of the family of nucleic acids of relatively high
PT significance -
XX
PS Disclosure; Page 120; 166pp; English.
XX
CC The present invention describes a collection of binding groups for a
CC family of nucleic acids comprising members of relative high and relative
CC low significance, where the binding groups are selected to be capable to
CC identify, alone or in combination, essentially all members of the family
CC of nucleic acids of relatively high significance. The collection of
CC binding groups is useful for typing of nucleic acid in a clinical sample,
CC by contacting the nucleic acid with the collection and determining
CC whether one or more binding groups bound to the nucleic acid of the
CC sample. This method is useful for determining whether the sample
CC comprises at least a part of a member of relatively high significance of
CC a family of nucleic acids. The collection of binding groups is useful for
CC diagnosing the severity of a disease caused by a pathogen containing a
CC member of a family of nucleic acids. ABL88779 to ABL89321 represent
CC oligonucleotide sequences used in the exemplification of the present
CC invention.
XX
SQ Sequence 18 BP; 3 A; 2 C; 8 G; 5 T; 0 other;

Query Match 69.0%; Score 13.8; DB 24; Length 18;
Best Local Similarity 88.2%; Pred. No. 2.3e+03;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AACATCTGTGGGGTTGG 17
||||||||| ||| |||
Db 1 AACATCTGTTGGGGTGG 17

RESULT 8
ABL89265
ID ABL89265 standard; DNA; 18 BP.
XX
AC ABL89265;
XX
DT 22-MAY-2002 (first entry)
XX
DE HIV-1 related binding molecule oligonucleotide sequence SEQ ID NO:487.
XX
KW Binding molecule; HIV-1; human immunodeficiency virus type 1;
KW reverse transcriptase; binding group; ss.
XX
OS Human immunodeficiency virus type 1.
OS Synthetic.
XX
PN EP1174518-A1.
XX
PD 23-JAN-2002.
XX
PF 20-JUL-2000; 2000EP-0202611.
XX
PR 20-JUL-2000; 2000EP-0202611.

XX
PA (AMST-) AMSTERDAM SUPPORT DIAGNOSTICS BV.
XX
PI Loukachov VV, Van Gemen B, Goudsmit J;
XX
DR WPI; 2002-156696/21.
XX
PT Collection of binding groups for determining or typing samples,
PT especially clinical samples, has groups capable to identify essentially
PT all members of the family of nucleic acids of relatively high
PT significance -
XX
PS Disclosure; Page 125; 166pp; English.
XX
CC The present invention describes a collection of binding groups for a
CC family of nucleic acids comprising members of relative high and relative
CC low significance, where the binding groups are selected to be capable to
CC identify, alone or in combination, essentially all members of the family
CC of nucleic acids of relatively high significance. The collection of
CC binding groups is useful for typing of nucleic acid in a clinical sample,
CC by contacting the nucleic acid with the collection and determining
CC whether one or more binding groups bound to the nucleic acid of the
CC sample. This method is useful for determining whether the sample
CC comprises at least a part of a member of relatively high significance of
CC a family of nucleic acids. The collection of binding groups is useful for
CC diagnosing the severity of a disease caused by a pathogen containing a
CC member of a family of nucleic acids. ABL88779 to ABL89321 represent
CC oligonucleotide sequences used in the exemplification of the present
CC invention.
XX
SQ Sequence 18 BP; 4 A; 2 C; 6 G; 6 T; 0 other;

Query Match 69.0%; Score 13.8; DB 24; Length 18;
Best Local Similarity 88.2%; Pred. No. 2.3e+03;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AACATCTGTGGGGTTGG 17
||||||||| | |||||
Db 1 AACATCTGTTGAGTTGG 17

RESULT 9
AAS05663
ID AAS05663 standard; DNA; 20 BP.
XX
AC AAS05663;
XX
DT 26-SEP-2001 (first entry)
XX
DE Rat NCAM reverse PCR primer.
XX
KW Nestin; mammal; pancreatic stem cell; cytokeratin-19; diabetes mellitus;
KW liver disease; progenitor cell; cell differentiation; insulin; ss; rat;
KW islet of Langerhans; islet ductal cell; PCR primer; NCAM.
XX
OS Rattus sp.
XX
PN WO200139784-A1.
XX
PD 07-JUN-2001.
XX
PF 06-DEC-2000; 2000WO-US33031.
XX
PR 06-DEC-1999; 99US-0169082.
PR 28-JUN-2000; 2000US-0215109.
PR 06-OCT-2000; 2000US-0238880.
XX
PA (GEHO) GEN HOSPITAL CORP.
XX
PI Abraham EJ, Faustman D, Habener JL, Vallejo M, Zulewski H;
XX
DR WPI; 2001-408256/43.
XX
PT Treating diabetes mellitus or liver disease, comprises isolating a
PT nestin-positive pancreatic stem cell from a pancreatic islet of a
PT donor, and transferring the stem cell into the patient -
XX
PS Example 6; Page 49; 102pp; English.
XX
CC The sequence represents a PCR primer used in the identification of
CC nestin-positive rat stem cells. Nestin has been identified as a molecular
CC marker for pancreatic stem cells, while cytokeratin-19 serves as a marker
CC for a class of islet ductal cells. Mammalian pancreatic stem cells may be
CC used in the treatment of diabetes mellitus and other disorders such as
CC liver disease, and in transplantation of stem cells, progenitor cells or
CC differentiated cells into a mammal for the production of hormones, such
CC as insulin. They are also useful for identifying, localising and
CC isolating pancreatic stem cells. A nestin-positive pancreatic stem cell
CC is isolated from a donor pancreatic islet of Langerhans, and then
CC transferred into a patient with diabetes, where it differentiates into an
CC insulin-producing cell. The stem cell can also be expanded ex vivo to
CC produce a progenitor cell, prior to its transfer into the patient, where
CC it differentiates into an insulin-producing beta cell. Alternatively, the
CC progenitor cell can be differentiated in culture to form pseudo-islet
CC like aggregates that are then transferred.
XX
SQ Sequence 20 BP; 3 A; 3 C; 7 G; 7 T; 0 other;

Query Match 69.0%; Score 13.8; DB 22; Length 20;
Best Local Similarity 88.2%; Pred. No. 2.4e+03;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 AACATCTGTGGGGTTGG 17
||| ||||||||||||
Db 4 AACTCCTGTGGGGTTGG 20

RESULT 10
ABL50359/c
ID ABL50359 standard; DNA; 23 BP.
XX
AC ABL50359;
XX
DT 13-JUN-2002 (first entry)
XX
DE Human cancer cell growth inhibiting protein PP5656 PCR primer #2.
XX
KW Human; cancer cell growth inhibitor; cancer; cytostatic; PCR primer; ss.
XX
OS Homo sapiens.
XX
PN CN1324820-A.
XX
PD 05-DEC-2001.
XX
PF 18-MAY-2000; 2000CN-0115745.
XX
PR 18-MAY-2000; 2000CN-0115745.
XX
PA (SHAN-) SHANGHAI CITY INST ONCOLOGY.
XX
PI Gu J, Yang S;
XX
DR WPI; 2002-281647/33.
XX
PT Novel human protein, and the polynucleotide that encodes it, for
PT inhibiting cancer cell growth -
XX
PS Example 2; Page 11 (Disclosure); 33pp; Chinese.
XX
CC ABL50333 to ABL50341 encode the human cancer cell growth inhibitory
CC proteins given in ABB06778 to ABB06786. The human cancer cell growth
CC inhibitory proteins and their encoding polynucleotide sequences have
CC cytostatic activity. They can be used for inhibiting cancer. The
CC present sequence represents a PCR primer for a human cancer cell
CC growth inhibitory protein, which is used in an example from the

```
CC   present invention.
XX
SQ   Sequence 23 BP; 7 A; 11 C; 3 G; 2 T; 0 other;

  Query Match             69.0%;  Score 13.8;  DB 24;  Length 23;
  Best Local Similarity   88.2%;  Pred. No. 2.4e+03;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 AACATCTGTGGGGTTGG 17
           | ||||||||||||  ||
Db      22 AGCATCTGTGGGGTGGG 6


RESULT 11
AAZ24764/c
ID   AAZ24764 standard; DNA; 32 BP.
XX
AC   AAZ24764;
XX
DT   11-FEB-2000  (first entry)
XX
DE   Human CCR5 receptor DNA specific primer.
XX
KW   Prostaglandin; PG; E2EP3 receptor; E2EP2 receptor; CCR-5; human; HIV;
KW   chemokine receptor; ss2 adrenergic receptor; small G-protein rho;
KW   renal outer medullary potassium ion channel protein; ion-channel protein;
KW   lambda phage repressor protein; G-protein coupled receptor; bacteria; ss;
KW   biochemical; vaccine; immunohistochemical; orphan receptor; PCR primer.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
PN   WO9953033-A1.
XX
PD   21-OCT-1999.
XX
PF   16-APR-1999;   99WO-US08214.
XX
PR   16-APR-1998;   98US-0081989.
XX
PA   (UYVA-) UNIV VANDERBILT.
XX
PI   Breyer RM,  Ma L,  Kennedy C;
XX
DR   WPI; 1999-620416/53.
XX
PT   New nucleic acid constructs for high level expression of eukaryotic
PT   proteins in bacteria, for producing e.g. chemokine receptor CCR-5 for
PT   preventing HIV infection  -
XX
PS   Example 3; Page 28; 81pp; English.
XX
CC   The invention provides isolated nucleic acid sequences that encode
CC   rabbit prostaglandin (PG) E2EP3 receptor, human PG E2EP2 receptor, human
CC   chemokine receptor CCR-5, human ss2 adrenergic receptor, rat renal outer
CC   medullary potassium ion channel protein or human small G-protein rho,
CC   together with deduced protein sequences. Also provided is a method for
CC   the production of eukaryotic proteins by culturing bacteria transformed
CC   with vectors containing the above nucleic acid sequences or a nucleic
CC   acid (I) that comprises: (i) first sequence that encodes either a
CC   sequence comprising at least three positively charged amino acids, or a
CC   DNA-binding protein, or a lambda phage repressor protein, placed
CC   upstream of, and in frame with, (ii) a sequence encoding a protein. (I)
CC   are used for recombinant production of eukaryotic proteins, particularly
CC   membrane proteins, G-protein coupled receptors or ion-channel proteins,
CC   in bacteria. These proteins are useful for biochemical or structural
CC   studies; as therapeutic agents; in diagnostic and screening assays and
CC   as antigens for use in vaccines, and for raising antibodies that are
CC   useful as immunohistochemical markers, e.g. for orphan receptors or ion
CC   channels. Antibodies raised against the chemokine receptor CCR-5 can be
CC   used (when administered as antiserum or generated in vivo) to prevent
CC   entry of human immune deficiency virus (HIV) into cells. Sequences
CC   AAZ24763 -64 represent primers for amplifying the human CCR5 DNA.
XX
SQ   Sequence 32 BP; 7 A; 12 C; 7 G; 6 T; 0 other;

  Query Match             69.0%;  Score 13.8;  DB 20;  Length 32;
  Best Local Similarity   88.2%;  Pred. No. 2.5e+03;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 AACATCTGTGGGGTTGG 17
           || |||||||||  ||||
Db      28 AATATCTGTGGGCTTGG 12


RESULT 12
ABI97596/c
ID   ABI97596 standard; DNA; 33 BP.
XX
AC   ABI97596;
XX
DT   18-FEB-2002  (first entry)
XX
DE   Endogenous human GPCR 3' primer SEQ ID NO: 48.
XX
KW   Human; G protein-coupled receptor; GPCR; non-endogenous; mutant;
KW   constitutively activated GPCR; agonist; disease; PCR primer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200177172-A2.
XX
PD   18-OCT-2001.
XX
PF   05-APR-2001; 2001WO-US11098.
XX
PR   07-APR-2000; 2000US-195747P.
XX
PA   (AREN-) ARENA PHARM INC.
XX
PI   Lehmann-Bruinsma K,  Liaw CW,  Lin I;
XX
DR   WPI; 2001-648759/74.
XX
PT   Identifying agonists of G protein-coupled receptors (GPCRs) for use in
PT   disease treatment, comprises contacting candidate compounds with
PT   versions of GPCRs -
XX
PS   Example 1; Page 29; 394pp; English.
XX
CC   The invention relates to G protein-coupled receptors (GPCRs) for which
CC   the endogenous ligand has been identified. Non-endogenous
CC   constitutively activated versions of known GPCRs are used in the
CC   invention for the direct identification of candidate compounds as
CC   receptor agonists, inverse agonists or partial agonists. Such
CC   agonists are useful as therapeutic agents for diseases or disorders
CC   associated with GPCRs. The present sequence is a primer used to
CC   prepare an endogenous version of a known GPCR in an example
CC   illustrating the invention.
XX
SQ   Sequence 33 BP; 7 A; 13 C; 5 G; 8 T; 0 other;

  Query Match             69.0%;  Score 13.8;  DB 23;  Length 33;
  Best Local Similarity   88.2%;  Pred. No. 2.5e+03;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 AACATCTGTGGGGTTGG 17
           || |||||||||  ||||
Db      25 AATATCTGTGGGCTTGG 9


RESULT 13
AAL29283
ID   AAL29283 standard; DNA; 50 BP.
XX
AC   AAL29283;
```

```
XX
DT   24-JAN-2002  (first entry)
XX
DE   Human SNP oligonucleotide #2491.
XX
KW   Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
KW   neuroprotective; antimicrobial; gene therapy; vaccine; amylase; cancer;
KW   amyloid protein; angiopoietin; apoptosis related protein; cadherin;
KW   cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
KW   complement related protein; cytochrome; kinesin; cytokine; interferon;
KW   interleukin; G-protein coupled receptor; thioesterase; inflammation;
KW   multifactorial disease; autoimmune disease; infection;
KW   nervous system disease; ss.
XX
OS   Homo sapiens.
XX
PN   WO200147944-A2.
XX
PD   05-JUL-2001.
XX
PF   28-DEC-2000; 2000WO-US35498.
XX
PR   28-DEC-1999;   99US-0173419.
PR   27-DEC-2000; 2000US-0173419.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Shimkets RA,  Leach M;
XX
DR   WPI; 2001-465210/50.
XX
PT   Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
PT   oncogenes and histones, useful for diagnosing and treating, e.g.
PT   cancer, autoimmune diseases and infections -
XX
PS   Claim 1; Page 2096; 4143pp; English.
XX
CC   The present invention relates to oligonucleotides encoding polymorphic
CC   variants of proteins related to amylases, amyloid proteins, angiopoietin,
CC   apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,
CC   histones, kinases, colony stimulating factors, complement related
CC   proteins, cytochromes, kinesins, cytokines, interferons, interleukins,
CC   G-protein coupled receptors and thioesterases. The present sequence is
CC   one such oligonucleotide. The oligonucleotides and the peptides encoded
CC   by them may be used in the prevention, diagnosis and treatment of
CC   diseases associated with inappropriate expression of the proteins listed
CC   above. Disorders that may be prevented, diagnosed and/or treated include
CC   multifactorial diseases with a genetic component, such as autoimmune
CC   diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,
CC   systemic lupus erythromatosus and Grave's disease), inflammation, cancer
CC   (e.g. cancers of the bladder, brain, breast, colon and kidney,
CC   leukaemia), diseases of the nervous system and an infection of pathogenic
CC   organisms.
XX
SQ   Sequence 50 BP; 6 A; 18 C; 17 G; 9 T; 0 other;

  Query Match             69.0%;  Score 13.8;  DB 22;  Length 50;
  Best Local Similarity   88.2%;  Pred. No. 2.6e+03;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 AACATCTGTGGGGTTGG 17
           |||||| ||||||| ||
Db       8 AACATCCGTGGGGTGGG 24


RESULT 14
AAL29893
ID   AAL29893 standard; DNA; 51 BP.
XX
AC   AAL29893;
XX
DT   24-JAN-2002  (first entry)
XX
DE   Human SNP oligonucleotide #3101.
XX
KW   Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
KW   neuroprotective; antimicrobial; gene therapy; vaccine; amylase; cancer;
KW   amyloid protein; angiopoietin; apoptosis related protein; cadherin;
KW   cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
KW   complement related protein; cytochrome; kinesin; cytokine; interferon;
KW   interleukin; G-protein coupled receptor; thioesterase; inflammation;
KW   multifactorial disease; autoimmune disease; infection;
KW   nervous system disease; ss.
XX
OS   Homo sapiens.
XX
PN   WO200147944-A2.
XX
PD   05-JUL-2001.
XX
PF   28-DEC-2000; 2000WO-US35498.
XX
PR   28-DEC-1999;   99US-0173419.
PR   27-DEC-2000; 2000US-0173419.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Shimkets RA,  Leach M;
XX
DR   WPI; 2001-465210/50.
XX
PT   Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
PT   oncogenes and histones, useful for diagnosing and treating, e.g.
PT   cancer, autoimmune diseases and infections -
XX
PS   Claim 1; Page 2274; 4143pp; English.
XX
CC   The present invention relates to oligonucleotides encoding polymorphic
CC   variants of proteins related to amylases, amyloid proteins, angiopoietin,
CC   apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,
CC   histones, kinases, colony stimulating factors, complement related
CC   proteins, cytochromes, kinesins, cytokines, interferons, interleukins,
CC   G-protein coupled receptors and thioesterases. The present sequence is
CC   one such oligonucleotide. The oligonucleotides and the peptides encoded
CC   by them may be used in the prevention, diagnosis and treatment of
CC   diseases associated with inappropriate expression of the proteins listed
CC   above. Disorders that may be prevented, diagnosed and/or treated include
CC   multifactorial diseases with a genetic component, such as autoimmune
CC   diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,
CC   systemic lupus erythromatosus and Grave's disease), inflammation, cancer
CC   (e.g. cancers of the bladder, brain, breast, colon and kidney,
CC   leukaemia), diseases of the nervous system and an infection of pathogenic
CC   organisms.
XX
SQ   Sequence 51 BP; 15 A; 13 C; 11 G; 12 T; 0 other;

  Query Match             69.0%;  Score 13.8;  DB 22;  Length 51;
  Best Local Similarity   88.2%;  Pred. No. 2.6e+03;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 AACATCTGTGGGGTTGG 17
           ||||||| ||||| |||
Db      21 AACATCTTTGGGGGTGG 37


RESULT 15
AAV23400
ID   AAV23400 standard; DNA; 73 BP.
XX
AC   AAV23400;
XX
DT   08-JUL-1998  (first entry)
XX
DE   Template extension molecule used in method of the invention.
XX
KW   PCR primer; vaccine classification; poliovirus type 2 vaccine;
```

KW attenuated poliovirus vaccine; neurovirulence; TEM;
KW template extension molecule; ss.
XX
OS Synthetic.
OS Poliovirus.
XX
PN US5728519-A.
XX
PD 17-MAR-1998.
XX
PF 21-DEC-1994; 94US-0361337.
XX
PR 21-DEC-1994; 94US-0361337.
PR 06-NOV-1990; 90US-0607742.
PR 18-MAY-1994; 94US-0246373.
XX
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Chumakov KM, Levenbook IS, Norwood LP, Roninson I;
XX
DR WPI; 1998-270433/24.
XX
PT Determining acceptability of poliovirus vaccines - based on mutation
PT reversion(s) and comparison to World Health Organisation standard
XX
PS Example; Column 12; 38pp; English.
XX
CC This sequence represents a template extension molecule used in the method
CC of the invention. The method is for classifying an unclassified live
CC poliovirus type 2 vaccine (attenuated by a G to A substitution at
CC nucleotide position 481) as having an acceptable or unacceptable level of
CC neurovirulence, and comprises, prior to vaccine administration: (a)
CC selectively amplifying a region (I) of the poliovirus genome containing
CC nucleotide position 481 using selectively mismatched primers to introduce
CC a site-specific mutation to create a restriction endonuclease (RE) site
CC which includes nucleotide position 481; (b) digesting an amount of (I)
CC with a RE that specifically cleaves the amplified sequences in revertant
CC viruses which contain an A to G reversion at nucleotide position 481; (c)
CC digesting an amount of (I) with a RE that specifically cleaves the
CC amplified sequences in non-revertant viruses which contain an A at
CC nucleotide position 481; (d) quantifying the percentage of revertant
CC viruses in the unclassified vaccine; and (e) comparing the percentage of
CC revertant viruses in the unclassified vaccine to the percentage of
CC revertant viruses in an accepted reference vaccine which can pass the
CC monkey neurovirulence test utilised by the World Health Organisation, an
CC unclassified vaccine with a higher percentage of A to G revertant viruses
CC than in the reference vaccine being classified as unacceptable and an
CC unclassified vaccine with an equal or lower percentage of A to G
CC revertant viruses than in the reference vaccine classified as acceptable.
CC The test can also be used to identify cells that are suitable for the
CC culture of attenuated polioviruses.
XX
SQ Sequence 73 BP; 19 A; 19 C; 20 G; 15 T; 0 other;

Query Match 69.0%; Score 13.8; DB 19; Length 73;
Best Local Similarity 88.2%; Pred. No. 2.7e+03;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy      1 AACATCTGTGGGGTTGG 17
          |||| || |||||||||
Db     40 AACAGCTCTGGGGTTGG 56
```

Search completed: November 23, 2002, 07:03:41
Job time : 97.1 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 06:42:25 ; Search time 16.8 Seconds
(without alignments)
450.869 Million cell updates/sec

Title: US-09-296-264-22
Perfect score: 20
Sequence: 1 aacatctgtggggttggtgt 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 335578 seqs, 189365133 residues

Total number of hits satisfying chosen parameters: 195614

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	13.8	69.0	20	9	US-09-963-875-12	Sequence 12, Appl
2	13.6	68.0	36	10	US-09-765-272-260	Sequence 260, App
3	13.2	66.0	90	10	US-09-864-761-28590	Sequence 28590, A
c 4	12.8	64.0	57	10	US-09-923-246-65	Sequence 65, Appl
c 5	12.8	64.0	61	10	US-09-725-285-6	Sequence 6, Appli
c 6	12.8	64.0	61	10	US-09-779-879A-6	Sequence 6, Appli
c 7	12.8	64.0	61	10	US-09-779-880A-6	Sequence 6, Appli
c 8	12.8	64.0	61	10	US-09-195-662A-6	Sequence 6, Appli
c 9	12.8	64.0	61	10	US-09-339-912A-6	Sequence 6, Appli
c 10	12.8	64.0	61	10	US-09-502-783A-6	Sequence 6, Appli
c 11	12.8	64.0	67	10	US-09-923-246-36	Sequence 36, Appl
c 12	12.8	64.0	67	10	US-09-825-561A-28	Sequence 28, Appl
13	12.6	63.0	32	10	US-09-770-693-20	Sequence 20, Appl
14	12.6	63.0	56	10	US-09-920-300A-1350	Sequence 1350, Ap
15	12.6	63.0	56	12	US-10-033-528-1350	Sequence 1350, Ap
16	12.6	63.0	97	10	US-09-864-761-23501	Sequence 23501, A
17	12.4	62.0	74	10	US-09-681-508-13	Sequence 13, Appl
c 18	12.2	61.0	18	10	US-09-932-679-4	Sequence 4, Appli
c 19	12.2	61.0	22	8	US-08-424-550B-145	Sequence 145, App
20	12.2	61.0	31	10	US-09-801-274-361	Sequence 361, App
c 21	12.2	61.0	31	10	US-09-801-274-671	Sequence 671, App
22	12.2	61.0	75	10	US-09-758-140-9	Sequence 9, Appli
23	12.2	61.0	75	10	US-09-972-599A-9	Sequence 9, Appli
24	12.2	61.0	75	10	US-09-972-599A-11	Sequence 11, Appl
25	12.2	61.0	75	10	US-09-972-599A-33	Sequence 33, Appl
26	12.2	61.0	90	10	US-09-972-599A-31	Sequence 31, Appl
c 27	12.2	61.0	96	10	US-09-960-352-7262	Sequence 7262, Ap
28	12	60.0	27	9	US-09-321-005A-8	Sequence 8, Appli
c 29	12	60.0	77	10	US-09-969-373-1390	Sequence 1390, Ap
c 30	12	60.0	89	10	US-09-864-761-22919	Sequence 22919, A
c 31	12	60.0	90	10	US-09-864-761-32534	Sequence 32534, A
c 32	11.8	59.0	29	10	US-09-725-285-4	Sequence 4, Appli
c 33	11.8	59.0	29	10	US-09-725-285-8	Sequence 8, Appli
c 34	11.8	59.0	29	10	US-09-779-879A-4	Sequence 4, Appli
c 35	11.8	59.0	29	10	US-09-779-879A-8	Sequence 8, Appli
c 36	11.8	59.0	29	10	US-09-779-880A-4	Sequence 4, Appli
c 37	11.8	59.0	29	10	US-09-779-880A-8	Sequence 8, Appli
c 38	11.8	59.0	29	10	US-09-195-662A-4	Sequence 4, Appli
c 39	11.8	59.0	29	10	US-09-195-662A-8	Sequence 8, Appli
c 40	11.8	59.0	29	10	US-09-339-912A-4	Sequence 4, Appli
c 41	11.8	59.0	29	10	US-09-339-912A-8	Sequence 8, Appli
c 42	11.8	59.0	29	10	US-09-502-783A-4	Sequence 4, Appli
c 43	11.8	59.0	29	10	US-09-502-783A-8	Sequence 8, Appli
c 44	11.8	59.0	51	10	US-09-814-865-1	Sequence 1, Appli
45	11.8	59.0	54	10	US-09-878-574-1788	Sequence 1788, Ap

ALIGNMENTS

RESULT 1
US-09-963-875-12
; Sequence 12, Application US/09963875
; Patent No. US20020164307A1
; GENERAL INFORMATION:
; APPLICANT: Massachusetts General Hospital
; TITLE OF INVENTION: Stem Cells of the Islets of Langerhans and Their Use in Treati
; TITLE OF INVENTION: Mellitus
; FILE REFERENCE: 17633/1235
; CURRENT APPLICATION NUMBER: US/09/963,875
; CURRENT FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: US60/169082
; PRIOR FILING DATE: 1999-12-06
; PRIOR APPLICATION NUMBER: US 60/215109
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: US 60/238880
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: US 09/731261
; PRIOR FILING DATE: 2000-12-06
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 12
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Oligonucleotide
US-09-963-875-12

Query Match 69.0%; Score 13.8; DB 9; Length 20;
Best Local Similarity 88.2%; Pred. No. 3.9e+02;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy       1 AACATCTGTGGGGTTGG 17
           ||| ||||||||||||
Db       4 AACTCCTGTGGGGTTGG 20
```

RESULT 2
US-09-765-272-260
; Sequence 260, Application US/09765272
; Patent No. US20020061545A1

```
;   GENERAL INFORMATION:
;        APPLICANT: Choi et. al.
;        TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
;        NUMBER OF SEQUENCES: 452
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Human Genome Sciences, Inc.
;             STREET: 9410 Key West Avenue
;             CITY: Rockville
;             STATE: Maryland
;             COUNTRY: USA
;             ZIP: 20850
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
;             COMPUTER: HP Vectra 486/33
;             OPERATING SYSTEM: MSDOS version 6.2
;             SOFTWARE: ASCII Text
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/765,272
;             FILING DATE: 22-Jan-2001
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 08/961,083
;             FILING DATE: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Brookes, A. Anders
;             REGISTRATION NUMBER: 36,373
;             REFERENCE/DOCKET NUMBER: PB340P2
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (301) 309-8504
;             TELEFAX: (301) 309-8512
;   INFORMATION FOR SEQ ID NO: 260:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 36 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: double
;             TOPOLOGY: linear
;        SEQUENCE DESCRIPTION: SEQ ID NO: 260:
US-09-765-272-260

  Query Match             68.0%;  Score 13.6;  DB 10;  Length 36;
  Best Local Similarity   80.0%;  Pred. No. 5.3e+02;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 AACATCTGTGGGGTTGGTGT 20
          | | ||||| || |||||||
Db      6 AGCTTCTGTAGGCTTGGTGT 25


RESULT 3
US-09-864-761-28590
; Sequence 28590, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 28590
;   LENGTH: 90
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AC018367.3
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.86
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.64
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.73
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.91
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.75
;   OTHER INFORMATION: NT HIT: AL161510.2, EVALUE 2.90e+00
;   OTHER INFORMATION: EST_HUMAN HIT: AI478339.1, EVALUE 1.10e+00
US-09-864-761-28590

  Query Match             66.0%;  Score 13.2;  DB 10;  Length 90;
  Best Local Similarity   83.3%;  Pred. No. 9.5e+02;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 ACATCTGTGGGGTTGGTG 19
          |||| ||||  |||||||
Db     73 ACATGTGTGATGTTGGTG 90


RESULT 4
US-09-923-246-65/c
; Sequence 65, Application US/09923246
; Patent No. US20020128446A1
; GENERAL INFORMATION:
;  APPLICANT; No. US20020128446A1ak, Julia E.
;  APPLICANT:  Presnell, Scott R.
;  APPLICANT:  Sprecher, Cindy A.
;  APPLICANT:  Foster, Donald C.
;  APPLICANT:  Holly, Richard D.
;  APPLICANT:  Gross, Jane A.
;  APPLICANT:  Johnston, Janet V.
;  APPLICANT:  Nelson, Andrew J.
;  APPLICANT:  Dillon, Stacey R.
;  APPLICANT:  Hammond, Angela K.
;  TITLE OF INVENTION: NOVEL CYTOKINE ZALPHA11 LIGAND
;  FILE REFERENCE: 99-16
;  CURRENT APPLICATION NUMBER: US/09/923,246
;  CURRENT FILING DATE:  2001-08-03
;  PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/522,217
;  PRIOR FILING DATE: EARLIER FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/123,904
```

```
;  PRIOR FILING DATE: EARLIER FILING DATE: 1999-03-11
;  PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/142,013
;  PRIOR FILING DATE: EARLIER FILING DATE: 1999-07-01
;  NUMBER OF SEQ ID NOS: 115
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 65
;   LENGTH: 57
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Oligonucleotide primer ZC22053
US-09-923-246-65

  Query Match            64.0%;  Score 12.8;  DB 10;  Length 57;
  Best Local Similarity  87.5%;  Pred. No. 1.4e+03;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 AACATCTGTGGGGTTG 16
          |||| || ||||||||
Db     19 AACAGCTCTGGGGTTG 4


RESULT 5
US-09-725-285-6/c
; Sequence 6, Application US/09725285
; Patent No. US20010000241A1
; GENERAL INFORMATION:
;  APPLICANT:   Li, Yi
;  APPLICANT:  Ruben, Steven, M.
;  TITLE OF INVENTION:   Antibodies to Human G-Protein Chemokine Receptor HDGNR10
;  TITLE OF INVENTION:  (CCR5 Receptor)
;  FILE REFERENCE:   1488.1150003
;  CURRENT APPLICATION NUMBER: US/09/725,285
;  CURRENT FILING DATE:  2000-11-29
;  PRIOR APPLICATION NUMBER:   09/339,912
;  PRIOR FILING DATE:   1999-06-25
;  PRIOR APPLICATION NUMBER:   09/195,662
;  PRIOR FILING DATE:   1998-11-18
;  PRIOR APPLICATION NUMBER:   08/466,343
;  PRIOR FILING DATE:   1995-06-06
;  NUMBER OF SEQ ID NOS:   9
;  SOFTWARE:   PatentIn version 3.0
; SEQ ID NO 6
;   LENGTH: 61
;   TYPE: DNA
;   ORGANISM: Oligonucleotide
US-09-725-285-6

  Query Match            64.0%;  Score 12.8;  DB 10;  Length 61;
  Best Local Similarity  87.5%;  Pred. No. 1.4e+03;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 AACATCTGTGGGGTTG 16
          || |||||||||  |||
Db     59 AATATCTGTGGGCTTG 44


RESULT 6
US-09-779-879A-6/c
; Sequence 6, Application US/09779879A
; Patent No. US20020048786A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen, Craig A.
;  APPLICANT:  Roschke, Viktor
;  APPLICANT:  Li, Yi
;  APPLICANT:  Ruben, Steven, M.
;  TITLE OF INVENTION: Human G-protein Chemokine Receptor (CCR5) HDGNR10
;  FILE REFERENCE: 1488.115000A
;  CURRENT APPLICATION NUMBER: US/09/779,879A
;  CURRENT FILING DATE:  2001-02-09
;  PRIOR APPLICATION NUMBER: US 60/181,258
;  PRIOR FILING DATE: 2000-02-09
;  PRIOR APPLICATION NUMBER: US 60/187,999
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: US 60/234,336
;  PRIOR FILING DATE: 2000-09-22
;  NUMBER OF SEQ ID NOS: 58
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
;   LENGTH: 61
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: 3' Oligonucleotide primer for HDGNR10
US-09-779-879A-6

  Query Match            64.0%;  Score 12.8;  DB 10;  Length 61;
  Best Local Similarity  87.5%;  Pred. No. 1.4e+03;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 AACATCTGTGGGGTTG 16
          || |||||||||  |||
Db     59 AATATCTGTGGGCTTG 44


RESULT 7
US-09-779-880A-6/c
; Sequence 6, Application US/09779880A
; Patent No. US20020061834A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen, Craig A.
;  APPLICANT:  Roschke, Viktor
;  APPLICANT:  Li, Yi
;  APPLICANT:  Ruben, Steven, M.
;  TITLE OF INVENTION: Human G-protein Chemokine Receptor (CCR5) HDGNR10
;  FILE REFERENCE: 1488.115000C
;  CURRENT APPLICATION NUMBER: US/09/779,880A
;  CURRENT FILING DATE:  2001-02-09
;  PRIOR APPLICATION NUMBER: US 60/181,258
;  PRIOR FILING DATE: 2000-02-09
;  PRIOR APPLICATION NUMBER: US 60/187,999
;  PRIOR FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: US 60/234,336
;  PRIOR FILING DATE: 2000-09-22
;  NUMBER OF SEQ ID NOS: 58
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
;   LENGTH: 61
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: 3' Oligonucleotide primer for HDGNR10
US-09-779-880A-6

  Query Match            64.0%;  Score 12.8;  DB 10;  Length 61;
  Best Local Similarity  87.5%;  Pred. No. 1.4e+03;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 AACATCTGTGGGGTTG 16
          || |||||||||  |||
Db     59 AATATCTGTGGGCTTG 44


RESULT 8
US-09-195-662A-6/c
; Sequence 6, Application US/09195662A
; Patent No. US20020076745A1
; GENERAL INFORMATION:
;  APPLICANT:   Li, Yi
;  APPLICANT:  Ruben, Steven, M.
;  TITLE OF INVENTION:   Human G-Protein Chemokine Receptor HDGNR10 (CCR5 Receptor)
;  FILE REFERENCE:   1488.1150002
;  CURRENT APPLICATION NUMBER: US/09/195,662A
;  CURRENT FILING DATE:  1998-11-18
;  PRIOR APPLICATION NUMBER:   08/466,343
;  PRIOR FILING DATE:   1995-06-06
```

```
;  NUMBER OF SEQ ID NOS:   9
;  SOFTWARE:   PatentIn version 3.0
; SEQ ID NO 6
;   LENGTH: 61
;   TYPE: DNA
;   ORGANISM: Oligonucleotide
US-09-195-662A-6

  Query Match             64.0%;  Score 12.8;  DB 10;  Length 61;
  Best Local Similarity   87.5%;  Pred. No. 1.4e+03;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 AACATCTGTGGGGTTG 16
          || |||||||||  |||
Db     59 AATATCTGTGGGCTTG 44


RESULT 9
US-09-339-912A-6/c
; Sequence 6, Application US/09339912A
; Patent No. US20020099176A1
; GENERAL INFORMATION:
;  APPLICANT:   Li, Yi
;  APPLICANT:  Ruben, Steven, M.
;  TITLE OF INVENTION:   Antibodies to Human G-Protein Chemokine Receptor HDGNR10
;  TITLE OF INVENTION:  (CCR5 Receptor)
;  FILE REFERENCE:   1488.1150003
;  CURRENT APPLICATION NUMBER: US/09/339,912A
;  CURRENT FILING DATE:  1999-06-25
;  PRIOR APPLICATION NUMBER:   09/195,662
;  PRIOR FILING DATE:   1998-11-18
;  PRIOR APPLICATION NUMBER:   08/466,343
;  PRIOR FILING DATE:   1995-06-06
;  NUMBER OF SEQ ID NOS:   9
;  SOFTWARE:   PatentIn version 3.0
; SEQ ID NO 6
;   LENGTH: 61
;   TYPE: DNA
;   ORGANISM: Oligonucleotide
US-09-339-912A-6

  Query Match             64.0%;  Score 12.8;  DB 10;  Length 61;
  Best Local Similarity   87.5%;  Pred. No. 1.4e+03;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 AACATCTGTGGGGTTG 16
          || |||||||||  |||
Db     59 AATATCTGTGGGCTTG 44


RESULT 10
US-09-502-783A-6/c
; Sequence 6, Application US/09502783A
; Patent No. US20020132269A1
; GENERAL INFORMATION:
;  APPLICANT: Li, Yi
;  APPLICANT:  Ruben, Steven M.
;  TITLE OF INVENTION: Polynucleotides Encoding Human G-Protein Chemokine Receptor (CCR5
;  TITLE OF INVENTION:  HDGNR10
;  FILE REFERENCE: 1488.1150006
;  CURRENT APPLICATION NUMBER: US/09/502,783A
;  CURRENT FILING DATE:  2001-08-23
;  PRIOR APPLICATION NUMBER: 08/466,343
;  PRIOR FILING DATE: 1995-06-06
;  NUMBER OF SEQ ID NOS: 9
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
;   LENGTH: 61
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Oligonucleotide
US-09-502-783A-6

  Query Match             64.0%;  Score 12.8;  DB 10;  Length 61;
  Best Local Similarity   87.5%;  Pred. No. 1.4e+03;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 AACATCTGTGGGGTTG 16
          || |||||||||  |||
Db     59 AATATCTGTGGGCTTG 44


RESULT 11
US-09-923-246-36/c
; Sequence 36, Application US/09923246
; Patent No. US20020128446A1
; GENERAL INFORMATION:
;  APPLICANT: No. US20020128446A1ak, Julia E.
;  APPLICANT:  Presnell, Scott R.
;  APPLICANT:  Sprecher, Cindy A.
;  APPLICANT:  Foster, Donald C.
;  APPLICANT:  Holly, Richard D.
;  APPLICANT:  Gross, Jane A.
;  APPLICANT:  Johnston, Janet V.
;  APPLICANT:  Nelson, Andrew J.
;  APPLICANT:  Dillon, Stacey R.
;  APPLICANT:  Hammond, Angela K.
;  TITLE OF INVENTION: NOVEL CYTOKINE ZALPHA11 LIGAND
;  FILE REFERENCE: 99-16
;  CURRENT APPLICATION NUMBER: US/09/923,246
;  CURRENT FILING DATE:  2001-08-03
;  PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/522,217
;  PRIOR FILING DATE: EARLIER FILING DATE: 2000-03-09
;  PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/123,904
;  PRIOR FILING DATE: EARLIER FILING DATE: 1999-03-11
;  PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/142,013
;  PRIOR FILING DATE: EARLIER FILING DATE: 1999-07-01
;  NUMBER OF SEQ ID NOS: 115
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 36
;   LENGTH: 67
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Oligonucleotide primer spanning  the 3' end of Fc4
;   OTHER INFORMATION: and  the vector flanking region
US-09-923-246-36

  Query Match             64.0%;  Score 12.8;  DB 10;  Length 67;
  Best Local Similarity   87.5%;  Pred. No. 1.4e+03;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 AACATCTGTGGGGTTG 16
          |||| || ||||||||
Db     36 AACAGCTCTGGGGTTG 21


RESULT 12
US-09-825-561A-28/c
; Sequence 28, Application US/09825561A
; Patent No. US20020137677A1
; GENERAL INFORMATION:
;  APPLICANT: Sprecher, Cindy A.
;  APPLICANT:  No. US20020137677A1ak, Julia E.
;  APPLICANT:  West, James W.
;  APPLICANT:  Presnell, Scott R.
;  APPLICANT:  Holly, Richard D.
;  APPLICANT:  Nelson, Andrew J.
;  TITLE OF INVENTION: SOLUBLE ZALPHA11 CYTOKINE RECEPTORS
;  FILE REFERENCE: 00-22
;  CURRENT APPLICATION NUMBER: US/09/825,561A
;  CURRENT FILING DATE:  2000-04-05
;  PRIOR APPLICATION NUMBER: US 60/194,731
;  PRIOR FILING DATE: 2000-04-05
;  PRIOR APPLICATION NUMBER: US 60/222,121
```

```
;  PRIOR FILING DATE: 2000-07-28
;  NUMBER OF SEQ ID NOS: 86
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 28
;    LENGTH: 67
;    TYPE: DNA
;    ORGANISM: Artificial Sequence
;    FEATURE:
;    OTHER INFORMATION: Oligonucleotide primer spanning the 3' end of Fc4
;    OTHER INFORMATION: and the vector flanking region
US-09-825-561A-28

  Query Match            64.0%;  Score 12.8;  DB 10;  Length 67;
  Best Local Similarity  87.5%;  Pred. No. 1.4e+03;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 AACATCTGTGGGGTTG 16
          |||| || ||||||||
Db     36 AACAGCTCTGGGGTTG 21


RESULT 13
US-09-770-693-20
; Sequence 20, Application US/09770693
; Patent No. US20020069434A1
; GENERAL INFORMATION:
;  APPLICANT: Beer, Steven V.
;  APPLICANT:  Bauer, David W.
;  TITLE OF INVENTION: OOMYCETE-RESISTANT TRANSGENIC PLANTS BY VIRTUE OF
;  TITLE OF INVENTION:  PATHOGEN-INDUCED EXPRESSION OF A HETEROLOGOUS
;  TITLE OF INVENTION:  HYPERSENSITIVE RESPONSE ELICITOR
;  FILE REFERENCE: 19603/2501
;  CURRENT APPLICATION NUMBER: US/09/770,693
;  CURRENT FILING DATE:  2001-01-26
;  PRIOR APPLICATION NUMBER: 60/178,565
;  PRIOR FILING DATE: 2000-01-26
;  NUMBER OF SEQ ID NOS: 26
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 20
;    LENGTH: 32
;    TYPE: DNA
;    ORGANISM: Artificial Sequence
;    FEATURE:
;    OTHER INFORMATION: Description of Artificial Sequence:  primer
US-09-770-693-20

  Query Match            63.0%;  Score 12.6;  DB 10;  Length 32;
  Best Local Similarity  78.9%;  Pred. No. 1.6e+03;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      2 ACATCTGTGGGGTTGGTGT 20
          | || |||| ||||||| |
Db      8 AGATATGTGTGGTTGGTCT 26


RESULT 14
US-09-920-300A-1350
; Sequence 1350, Application US/09920300A
; Patent No. US20020136728A1
; GENERAL INFORMATION:
;  APPLICANT: King, Gordon E.
;  APPLICANT:  Meagher, Madeleine Joy
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Secrist, Heather
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF COLON CANCER
;  FILE REFERENCE: 210121.547
;  CURRENT APPLICATION NUMBER: US/09/920,300A
;  CURRENT FILING DATE:  2001-07-31
;  NUMBER OF SEQ ID NOS: 1789
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1350
;    LENGTH: 56
;    TYPE: DNA
;    ORGANISM: Homo sapiens
US-09-920-300A-1350

  Query Match            63.0%;  Score 12.6;  DB 10;  Length 56;
  Best Local Similarity  78.9%;  Pred. No. 1.7e+03;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      2 ACATCTGTGGGGTTGGTGT 20
          | |||||||||| ||  ||
Db     34 AGATCTGTGGGGCTGTGGT 52


RESULT 15
US-10-033-528-1350
; Sequence 1350, Application US/10033528
; Patent No. US20020131971A1
; GENERAL INFORMATION:
;  APPLICANT: King, Gordon E.
;  APPLICANT:  Meagher, Madeleine Joy
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  Secrist, Heather
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;  TITLE OF INVENTION:  AND DIAGNOSIS OF COLON CANCER
;  FILE REFERENCE: 210121.547C1
;  CURRENT APPLICATION NUMBER: US/10/033,528
;  CURRENT FILING DATE:  2001-12-26
;  NUMBER OF SEQ ID NOS: 1896
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1350
;    LENGTH: 56
;    TYPE: DNA
;    ORGANISM: Homo sapiens
US-10-033-528-1350

  Query Match            63.0%;  Score 12.6;  DB 12;  Length 56;
  Best Local Similarity  78.9%;  Pred. No. 1.7e+03;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      2 ACATCTGTGGGGTTGGTGT 20
          | |||||||||| ||  ||
Db     34 AGATCTGTGGGGCTGTGGT 52


Search completed: November 23, 2002, 07:10:39
Job time : 17.8 secs
```

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 26, 2002, 04:16:10 ; Search time 798.5 Seconds
(without alignments)
405.648 Million cell updates/sec

Title: US-09-296-264-22
Perfect score: 20
Sequence: 1 aacatctgtggggttggtgt 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 357874

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
c	1	15.8	79.0	81	14	Z20220	Z20220 HSAAABQEG S
	2	15.4	77.0	55	17	B44192	B44192 HS-1059-A1-
	3	14.8	74.0	66	17	BH000352	BH000352 2M0288G10
c	4	14.4	72.0	86	12	BG652375	BG652375 sad66c04.
c	5	14.2	71.0	46	17	TA72E08Q	AL457957 T. brucei
c	6	14.2	71.0	85	17	TA143C10Q	AL467016 T. brucei
c	7	14.2	71.0	93	17	TA158G07Q	AL471982 T. brucei
c	8	14.2	71.0	94	17	TA102B02Q	AL460458 T. brucei
c	9	14	70.0	73	17	AZ597771	AZ597771 1M0411E14
c	10	13.8	69.0	98	13	BJ077550	BJ077550 BJ077550
c	11	13.6	68.0	40	17	BH847656	BH847656 SALK_0551
c	12	13.6	68.0	53	17	TA238E05Q	AL481309 T. brucei
	13	13.6	68.0	68	17	AZ791858	AZ791858 2M0041L21
	14	13.6	68.0	78	17	AZ377417	AZ377417 1M0131F14
	15	13.6	68.0	85	17	AZ626002	AZ626002 1M0466D03
c	16	13.6	68.0	93	9	AI188904	AI188904 qd36d02.x
	17	13.6	68.0	94	9	AI625266	AI625266 ts42g10.x
	18	13.6	68.0	94	9	AI672672	AI672672 we57c07.x
	19	13.6	68.0	95	9	AI299671	AI299671 qn12d04.x
	20	13.6	68.0	96	9	AI289192	AI289192 qn25h09.x
	21	13.6	68.0	96	10	AW166296	AW166296 xn50h06.x
	22	13.6	68.0	96	12	BG056332	BG056332 naf67b03.
	23	13.6	68.0	97	10	AW468562	AW468562 he41a10.x
	24	13.6	68.0	98	9	AI082481	AI082481 os71g12.x
	25	13.6	68.0	99	10	AW468637	AW468637 he42g04.x
	26	13.6	68.0	99	10	AW872960	AW872960 hq20e02.x
	27	13.4	67.0	48	17	AZ946920	AZ946920 2M0208L19
c	28	13.4	67.0	97	9	AA278432	AA278432 zs81b11.r
	29	13.4	67.0	100	10	BE166272	BE166272 MR3-HT049
c	30	13.4	67.0	100	12	BF745306	BF745306 CM4-BT085
	31	13.2	66.0	34	17	AZ393795	AZ393795 1M0157G05
	32	13.2	66.0	61	12	BG561340	BG561340 EtESTed83
	33	13.2	66.0	70	17	AL764786	AL764786 Arabidops
	34	13.2	66.0	71	17	AZ613659	AZ613659 1M0442C11
	35	13.2	66.0	73	9	AA782241	AA782241 ai31d03.s
	36	13.2	66.0	82	17	AZ806699	AZ806699 2M0068L20
	37	13.2	66.0	88	17	AZ875397	AZ875397 2M0189E08
	38	13.2	66.0	91	17	AZ400080	AZ400080 1M0166P07
	39	13.2	66.0	94	9	AA688536	AA688536 vs18g09.r
c	40	13.2	66.0	94	9	AA215033	AA215033 mu84a08.r
c	41	13	65.0	77	14	BQ393045	BQ393045 NISC_mq28
c	42	13	65.0	94	9	AA426845	AA426845 vf25g10.r
	43	12.8	64.0	56	14	U77317	U77317 HSU77317 Hu
c	44	12.8	64.0	56	14	U77326	U77326 HSU77326 Hu
c	45	12.8	64.0	80	10	AV851794	AV851794 AV851794

ALIGNMENTS

RESULT 1
Z20220/c
LOCUS Z20220 81 bp mRNA linear EST 07-FEB-1995
DEFINITION HSAAABQEG S, Human foetal Adrenals tissue Homo sapiens cDNA, mRNA sequence.
ACCESSION Z20220
VERSION Z20220.1 GI:26965
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 81)
AUTHORS MRC Human Genome Mapping Project Resource Centre.
TITLE The UK-HGMP cDNA program
JOURNAL Unpublished (1993)
COMMENT Contact: MRC Human Genome Mapping Project Resource Centre
Clinical Research Centre
Watford Road,Harrow,Middlesex HA1 3UJ,U.K.
Email: biohelp@hgmp.mrc.ac.uk
single read.
FEATURES Location/Qualifiers
source 1. .81
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="S, Human foetal Adrenals tissue"
/note="Vector: Bluescript; clone_library=S, Human foetal Adrenals tissue; cloning vector is Bluescript. "
BASE COUNT 28 a 22 c 13 g 18 t

```
ORIGIN

  Query Match            79.0%;  Score 15.8;  DB 14;  Length 81;
  Best Local Similarity  89.5%;  Pred. No. 3e+03;
  Matches   17;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       2 ACATCTGTGGGGTTGGTGT 20
           |||| | ||||||||||||
Db      42 ACATTTTTGGGGTTGGTGT 24


RESULT 2
B44192
LOCUS       B44192                   55 bp    DNA     linear   GSS 21-OCT-1997
DEFINITION  HS-1059-A1-A02-MR.abi CIT Human Genomic Sperm Library C Homo
            sapiens genomic clone Plate=CT 781 Col=3 Row=A, DNA sequence.
ACCESSION   B44192
VERSION     B44192.1  GI:2549026
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 55)
  AUTHORS   Mahairas,G.G., Zackrone,K.D., Smith,T., Tipton,S., Schmidt,S.,
            Traicoff,R., Abajian,C., Blanchard,A., West,A. and Hood,L.E.
  TITLE     Construction of a Characterized Clone Resource for Genomic
            Sequencing: Generation and Preliminary Analysis of 20,000 Sequence
            Tagged Connectors
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Mahairas GG, Zackrone KD, Hood L
            University of Washington
            Seattle, WA 98195, USA
            Tel: (206) 616-8744
            Fax: (206) 685-7301
            Email: kzackron@u.washington.edu
            Sequence Tagged Connector
            Plate: CT 781  row: A  column: 3
            Class: BAC ends
            High quality sequence stop: 55.
FEATURES             Location/Qualifiers
     source          1. .55
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="Plate=CT 781 Col=3 Row=A"
                     /clone_lib="CIT Human Genomic Sperm Library C"
                     /sex="M"
                     /note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in
                     E-Coli DH10B"
BASE COUNT        7 a     10 c     17 g     21 t
ORIGIN

  Query Match            77.0%;  Score 15.4;  DB 17;  Length 55;
  Best Local Similarity  94.1%;  Pred. No. 4e+03;
  Matches   16;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       4 ATCTGTGGGGTTGGTGT 20
           | ||||||||||||||||
Db      34 AGCTGTGGGGTTGGTGT 50


RESULT 3
BH000352
LOCUS       BH000352                 66 bp    DNA     linear   GSS 27-APR-2001
DEFINITION  2M0288G10F Mouse 10kb plasmid UUGC2M library Mus musculus genomic
            clone UUGC2M0288G10 F, DNA sequence.
ACCESSION   BH000352
VERSION     BH000352.1  GI:13871578
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 66)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0288  row: G  column: 10
            Seq primer: CGTTGTAAAACGACGGCCAGT
            Class: plasmid ends
            High quality sequence stop: 66.
FEATURES             Location/Qualifiers
     source          1. .66
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC2M0288G10"
                     /clone_lib="Mouse 10kb plasmid UUGC2M library"
                     /sex="Female"
                     /lab_host="E. coli strain XL10-Gold, T1-resistant, F-"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (female) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
BASE COUNT        8 a     15 c     15 g     28 t
ORIGIN

  Query Match            74.0%;  Score 14.8;  DB 17;  Length 66;
  Best Local Similarity  88.9%;  Pred. No. 7.9e+03;
  Matches   16;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       3 CATCTGTGGGGTTGGTGT 20
           | |||||||||| ||||
Db      31 CTTCTGTGGGGTTAGTGT 48


RESULT 4
BG652375/c
LOCUS       BG652375                 86 bp    mRNA    linear   EST 29-NOV-2001
DEFINITION  sad66c04.y1 Gm-c1051 Glycine max cDNA clone GENOME SYSTEMS CLONE
            ID: Gm-c1051-5119 5', mRNA sequence.
ACCESSION   BG652375
VERSION     BG652375.1  GI:13789784
KEYWORDS    EST.
SOURCE      soybean.
  ORGANISM  Glycine max
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
```

```
            Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
            Glycine.
REFERENCE   1  (bases 1 to 86)
  AUTHORS   Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Coryell,V., Khanna
            ,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
            Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers
            ,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk
            ,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann
            ,R., Waterston,R. and Wilson,R.
  TITLE     Public Soybean EST Project
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Shoemaker R/Public Soybean EST Project
            Public Soybean EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available through: ResGen, Invitrogen Corp. 2130
            South Memorial Parkway Huntsville, AL 35801 For further information
            call: (800)-533-4363 or contact via email: ccu@resgen.com
            High quality sequence stop: 72.
FEATURES             Location/Qualifiers
     source          1. .86
                     /organism="Glycine max"
                     /db_xref="taxon:3847"
                     /clone="GENOME SYSTEMS CLONE ID: Gm-c1051-5119"
                     /clone_lib="Gm-c1051"
                     /tissue_type="floral meristematic mRNA"
                     /lab_host="DH10B"
                     /note="Vector: pBluescript II SK+; Site_1: EcoRI; Site_2:
                     XhoI; The cDNA library was constructed from floral
                     meristematic mRNA provided by Dr. Halina Knap of Clemson
                     University. Complementary DNA was synthesized from mRNA
                     using a primer consisting of a poly(dT) sequence with a
                     XhoI restriction site. EcoRI adapters were ligated to the
                     blunt-ended cDNA fragments followed by XhoI digestion. The
                     cDNA fragments were directionally cloned into the
                     EcoRI-XhoI restriction site of the pBluescript vector. The
                     ligated cDNA fragments were transformed into DH10B host
                     cells (GibcoBRL). This library was constructed in the
                     laboratory of Dr. Randy Shoemaker."
BASE COUNT       23 a      28 c      14 g      21 t
ORIGIN

  Query Match             72.0%;  Score 14.4;  DB 12;  Length 86;
  Best Local Similarity   93.8%;  Pred. No. 1.3e+04;
  Matches   15;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 AACATCTGTGGGGTTG 16
           ||||||||||| ||||
Db      33 AACATCTGTGGAGTTG 18


RESULT 5
TA72E08Q/c
LOCUS       TA72E08Q                  46 bp    DNA     linear   GSS 13-DEC-2000
DEFINITION  T. brucei sheared genomic DNA clone 72e08, reverse sequence,
            genomic survey sequence.
ACCESSION   AL457957
VERSION     AL457957.1  GI:11859789
KEYWORDS    GSS.
SOURCE      Trypanosoma brucei.
  ORGANISM  Trypanosoma brucei
            Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
            Trypanosoma.
REFERENCE   1  (bases 1 to 46)
  AUTHORS   Hall,N., Bowman,S., Lennard,N.J., Doggett,J., Atkin,R.,
            Chillingworth,C., Ormond,D., Harris,B., El-Sayed,N., Hou,L.,
            Melville,S.E., Rajandream,M.A. and Barrell,B.G.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-DEC-2000) Trypanosoma brucei genome sequencing
            project, Sanger Centre, The Wellcome Trust Genome Campus, Hinxton,
            Cambridge CB10 1SA, E-mail: barrell@sanger.ac.uk and
            nhl@sanger.ac.uk
COMMENT     Constructed at the Institute for Genomic Research (TIGR),
            Rockville, MD. Genomic DNA isolated from a cloned population of
            Trypanosoma brucei (TREU927/4 GUTat 10.1) was mechanically sheared
            to give a tight size distribution (
            4 kb).  The v + i method used for the library construction is
            described in detail in Smith, H. and Venter, J.C. (Making small
            insert libraries for whole genome shotgun sequencing projects.  In
            Genome Sequencing: A Practical Approach, eds. M. Vaudin and B.
            Barrell, Oxford University Press, 1999).
            Email: nelsayed@tigr.org
            Details of T. brucei sequencing at the Sanger Centre are available
            at http://www.sanger.ac.uk/Projects/T_brucei/.
FEATURES             Location/Qualifiers
     source          1. .46
                     /organism="Trypanosoma brucei"
                     /strain="TREU927"
                     /db_xref="taxon:5691"
                     /clone="72e08"
BASE COUNT       19 a      9 c      7 g      11 t
ORIGIN

  Query Match             71.0%;  Score 14.2;  DB 17;  Length 46;
  Best Local Similarity   84.2%;  Pred. No. 1.3e+04;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 AACATCTGTGGGGTTGGTG 19
           ||||||| || | ||||||
Db      21 AACATCTATGTGTTTGGTG 3


RESULT 6
TA143C10Q/c
LOCUS       TA143C10Q                 85 bp    DNA     linear   GSS 13-DEC-2000
DEFINITION  T. brucei sheared genomic DNA clone 143c10, reverse sequence,
            genomic survey sequence.
ACCESSION   AL467016
VERSION     AL467016.1  GI:11836371
KEYWORDS    GSS.
SOURCE      Trypanosoma brucei.
  ORGANISM  Trypanosoma brucei
            Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
            Trypanosoma.
REFERENCE   1  (bases 1 to 85)
  AUTHORS   Hall,N., Bowman,S., Lennard,N.J., Doggett,J., Atkin,R.,
            Chillingworth,C., Ormond,D., Harris,B., El-Sayed,N., Hou,L.,
            Melville,S.E., Rajandream,M.A. and Barrell,B.G.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-DEC-2000) Trypanosoma brucei genome sequencing
            project, Sanger Centre, The Wellcome Trust Genome Campus, Hinxton,
            Cambridge CB10 1SA, E-mail: barrell@sanger.ac.uk and
            nhl@sanger.ac.uk
COMMENT     Constructed at the Institute for Genomic Research (TIGR),
            Rockville, MD. Genomic DNA isolated from a cloned population of
            Trypanosoma brucei (TREU927/4 GUTat 10.1) was mechanically sheared
            to give a tight size distribution (
            4 kb).  The v + i method used for the library construction is
            described in detail in Smith, H. and Venter, J.C. (Making small
            insert libraries for whole genome shotgun sequencing projects.  In
            Genome Sequencing: A Practical Approach, eds. M. Vaudin and B.
            Barrell, Oxford University Press, 1999).
            Email: nelsayed@tigr.org
            Details of T. brucei sequencing at the Sanger Centre are available
            at http://www.sanger.ac.uk/Projects/T_brucei/.
FEATURES             Location/Qualifiers
     source          1. .85
                     /organism="Trypanosoma brucei"
                     /strain="TREU927"
                     /db_xref="taxon:5691"
                     /clone="143c10"
BASE COUNT       27 a      22 c      15 g      21 t
ORIGIN
```

```
  Query Match             71.0%;  Score 14.2;  DB 17;  Length 85;
  Best Local Similarity   84.2%;  Pred. No. 1.6e+04;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 AACATCTGTGGGGTTGGTG 19
          ||||||| || | ||||||
Db     83 AACATCTATGTGTTTGGTG 65


RESULT 7
TA158G07Q/c
LOCUS       TA158G07Q                 93 bp    DNA     linear   GSS 13-DEC-2000
DEFINITION  T. brucei sheared genomic DNA clone 158g07, reverse sequence,
            genomic survey sequence.
ACCESSION   AL471982
VERSION     AL471982.1  GI:11837336
KEYWORDS    GSS.
SOURCE      Trypanosoma brucei.
  ORGANISM  Trypanosoma brucei
            Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
            Trypanosoma.
REFERENCE   1  (bases 1 to 93)
  AUTHORS   Hall,N., Bowman,S., Lennard,N.J., Doggett,J., Atkin,R.,
            Chillingworth,C., Ormond,D., Harris,B., El-Sayed,N., Hou,L.,
            Melville,S.E., Rajandream,M.A. and Barrell,B.G.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-DEC-2000) Trypanosoma brucei genome sequencing
            project, Sanger Centre, The Wellcome Trust Genome Campus, Hinxton,
            Cambridge CB10 1SA, E-mail: barrell@sanger.ac.uk and
            nh1@sanger.ac.uk
COMMENT     Constructed at the Institute for Genomic Research (TIGR),
            Rockville, MD. Genomic DNA isolated from a cloned population of
            Trypanosoma brucei (TREU927/4 GUTat 10.1) was mechanically sheared
            to give a tight size distribution (
            4 kb).  The v + i method used for the library construction is
            described in detail in Smith, H. and Venter, J.C. (Making small
            insert libraries for whole genome shotgun sequencing projects.  In
            Genome Sequencing: A Practical Approach, eds. M. Vaudin and B.
            Barrell, Oxford University Press, 1999).
            Email: nelsayed@tigr.org
            Details of T. brucei sequencing at the Sanger Centre are available
            at http://www.sanger.ac.uk/Projects/T_brucei/.
FEATURES             Location/Qualifiers
     source          1. .93
                     /organism="Trypanosoma brucei"
                     /strain="TREU927"
                     /db_xref="taxon:5691"
                     /clone="158g07"
BASE COUNT       33 a     22 c     15 g     23 t
ORIGIN

  Query Match             71.0%;  Score 14.2;  DB 17;  Length 93;
  Best Local Similarity   84.2%;  Pred. No. 1.7e+04;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 AACATCTGTGGGGTTGGTG 19
          ||||||| || | ||||||
Db     83 AACATCTATGTGTTTGGTG 65


RESULT 8
TA102B02Q/c
LOCUS       TA102B02Q                 94 bp    DNA     linear   GSS 13-DEC-2000
DEFINITION  T. brucei sheared genomic DNA clone 102b02, reverse sequence,
            genomic survey sequence.
ACCESSION   AL460458
VERSION     AL460458.1  GI:11831208
KEYWORDS    GSS.
SOURCE      Trypanosoma brucei.
  ORGANISM  Trypanosoma brucei
            Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
            Trypanosoma.
REFERENCE   1  (bases 1 to 94)
  AUTHORS   Hall,N., Bowman,S., Lennard,N.J., Doggett,J., Atkin,R.,
            Chillingworth,C., Ormond,D., Harris,B., El-Sayed,N., Hou,L.,
            Melville,S.E., Rajandream,M.A. and Barrell,B.G.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-DEC-2000) Trypanosoma brucei genome sequencing
            project, Sanger Centre, The Wellcome Trust Genome Campus, Hinxton,
            Cambridge CB10 1SA, E-mail: barrell@sanger.ac.uk and
            nh1@sanger.ac.uk
COMMENT     Constructed at the Institute for Genomic Research (TIGR),
            Rockville, MD. Genomic DNA isolated from a cloned population of
            Trypanosoma brucei (TREU927/4 GUTat 10.1) was mechanically sheared
            to give a tight size distribution (
            4 kb).  The v + i method used for the library construction is
            described in detail in Smith, H. and Venter, J.C. (Making small
            insert libraries for whole genome shotgun sequencing projects.  In
            Genome Sequencing: A Practical Approach, eds. M. Vaudin and B.
            Barrell, Oxford University Press, 1999).
            Email: nelsayed@tigr.org
            Details of T. brucei sequencing at the Sanger Centre are available
            at http://www.sanger.ac.uk/Projects/T_brucei/.
FEATURES             Location/Qualifiers
     source          1. .94
                     /organism="Trypanosoma brucei"
                     /strain="TREU927"
                     /db_xref="taxon:5691"
                     /clone="102b02"
BASE COUNT       33 a     24 c     16 g     21 t
ORIGIN

  Query Match             71.0%;  Score 14.2;  DB 17;  Length 94;
  Best Local Similarity   84.2%;  Pred. No. 1.7e+04;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 AACATCTGTGGGGTTGGTG 19
          ||||||| || | ||||||
Db     85 AACATCTATGTGTTTGGTG 67


RESULT 9
AZ597771/c
LOCUS       AZ597771                  73 bp    DNA     linear   GSS 13-DEC-2000
DEFINITION  1M0411E14R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0411E14 R, DNA sequence.
ACCESSION   AZ597771
VERSION     AZ597771.1  GI:11719961
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 73)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0411  row: E  column: 14
            Seq primer: CACACAGGAAACAGCTATGACC
            Class: plasmid ends
            High quality sequence stop: 73.
FEATURES             Location/Qualifiers
```

```
     source          1. .73
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC1M0411E14"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
BASE COUNT       15 a     29 c     19 g     10 t
ORIGIN

  Query Match             70.0%;  Score 14;  DB 17;  Length 73;
  Best Local Similarity   100.0%;  Pred. No. 1.9e+04;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       6 CTGTGGGGTTGGTG 19
           ||||||||||||||
Db      59 CTGTGGGGTTGGTG 46


RESULT 10
BJ077550/c
LOCUS       BJ077550                98 bp    mRNA    linear   EST 11-DEC-2001
DEFINITION  BJ077550 NIBB Mochii normalized Xenopus tailbud library Xenopus
            laevis cDNA clone XL061f16 3', mRNA sequence.
ACCESSION   BJ077550
VERSION     BJ077550.1  GI:17522466
KEYWORDS    EST.
SOURCE      African clawed frog.
  ORGANISM  Xenopus laevis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus.
REFERENCE   1  (bases 1 to 98)
  AUTHORS   Kitayama,A., Terasaka,C., Mochii,M., Ueno,N., Shin-i,T. and Kohara
            ,Y.
  TITLE     Expressed genes in X. laevis embryo
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Tadasu Shin-i
            Center For Genetic Resource Information
            National Institute of Genetics
            1111 Yata, Mishima, Shizuoka 411-8540, Japan
            Tel: 81-559-81-6856
            Fax: 81-559-81-6855
            Email: tshini@genes.nig.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .98
                     /organism="Xenopus laevis"
                     /db_xref="taxon:8355"
                     /clone="XL061f16"
                     /clone_lib="NIBB Mochii normalized Xenopus tailbud
                     library"
                     /tissue_type="whole embryo"
                     /dev_stage="stage 25"
                     /note="Vector: pBSRN3; Site_1: NotI; Site_2: EcoRI; cDNAs
                     were oligo-dT primed and directionally cloned. Staging
                     according to  Nieuwkoop and Faber. Library is subtracted
                     and was constructed by N. Garrett  and A.M. Zorn,
                     (Wellcome/CRC Institute). "
BASE COUNT       40 a     35 c      6 g     15 t      2 others
ORIGIN

  Query Match             69.0%;  Score 13.8;  DB 13;  Length 98;
  Best Local Similarity   88.2%;  Pred. No. 2.6e+04;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       4 ATCTGTGGGGTTGGTGT 20
           ||||||| ||| |||||
Db      21 ATCTGTGTGGTGGGTGT 5


RESULT 11
BH847656/c
LOCUS       BH847656                40 bp    DNA     linear   GSS 13-JUN-2002
DEFINITION  SALK_055162.41.55.x Arabidopsis thaliana TDNA insertion lines
            Arabidopsis thaliana genomic clone SALK_055162.41.55.x, DNA
            sequence.
ACCESSION   BH847656
VERSION     BH847656.1  GI:21418527
KEYWORDS    GSS.
SOURCE      thale cress.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 40)
  AUTHORS   Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R., Gadrinab
            ,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L., Shinn,P.
            , Zimmerman,J. and Ecker,J.R.
  TITLE     A Sequence-Indexed Library of Insertion Mutations in the
            Arabidopsis Genome
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Joseph R. Ecker
            Salk Institute Genomic Analysis Laboratory (SIGnAL)
            The Salk Institute for Biological Studies
            10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
            Tel: 858 453 4100 x1752
            Fax: 858 558 6379
            Email: ecker@salk.edu
            This is single pass sequence recovered from the left border of
            TDNA. This sequence lies within 300 bases of the 5' end of
            At2g30945 and an annotated exon of At2g30950.
            Class: TDNA tagged.
FEATURES             Location/Qualifiers
     source          1. .40
                     /organism="Arabidopsis thaliana"
                     /strain="Columbia 0"
                     /db_xref="taxon:3702"
                     /clone="SALK_055162.41.55.x"
                     /clone_lib="Arabidopsis thaliana TDNA insertion lines"
                     /note="PCR was performed on Arabidopsis thaliana lines
                     each of which contains  one or more TDNA insertion
                     elements. The resultant fragment for each  line was
                     directly sequenced to determine the genomic sequence at
                     the  site of insertion. Details of the protocols used can
                     be found at  http://signal.salk.edu/tdna_protocols.html"
BASE COUNT       19 a      8 c      7 g      6 t
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 17;  Length 40;
  Best Local Similarity   80.0%;  Pred. No. 2.3e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 AACATCTGTGGGGTTGGTGT 20
           ||| |||||  ||||| |||
Db      30 AACCTCTGTTTGGTTGTTGT 11
```

```
RESULT 12
TA238E05Q/c
LOCUS       TA238E05Q                53 bp    DNA     linear   GSS 13-DEC-2000
DEFINITION  T. brucei sheared genomic DNA clone 238e05, reverse sequence,
            genomic survey sequence.
ACCESSION   AL481309
VERSION     AL481309.1  GI:11847003
KEYWORDS    GSS.
SOURCE      Trypanosoma brucei.
  ORGANISM  Trypanosoma brucei
            Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
            Trypanosoma.
REFERENCE   1  (bases 1 to 53)
  AUTHORS   Hall,N., Bowman,S., Lennard,N.J., Doggett,J., Atkin,R.,
            Chillingworth,C., Ormond,D., Harris,B., El-Sayed,N., Hou,L.,
            Melville,S.E., Rajandream,M.A. and Barrell,B.G.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-DEC-2000) Trypanosoma brucei genome sequencing
            project, Sanger Centre, The Wellcome Trust Genome Campus, Hinxton,
            Cambridge CB10 1SA, E-mail: barrell@sanger.ac.uk and
            nh1@sanger.ac.uk
COMMENT     Constructed at the Institute for Genomic Research (TIGR),
            Rockville, MD. Genomic DNA isolated from a cloned population of
            Trypanosoma brucei (TREU927/4 GUTat 10.1) was mechanically sheared
            to give a tight size distribution (
            4 kb).  The v + i method used for the library construction is
            described in detail in Smith, H. and Venter, J.C. (Making small
            insert libraries for whole genome shotgun sequencing projects.  In
            Genome Sequencing: A Practical Approach, eds. M. Vaudin and B.
            Barrell, Oxford University Press, 1999).
            Email: nelsayed@tigr.org
            Details of T. brucei sequencing at the Sanger Centre are available
            at http://www.sanger.ac.uk/Projects/T_brucei/.
FEATURES             Location/Qualifiers
     source          1. .53
                     /organism="Trypanosoma brucei"
                     /strain="TREU927"
                     /db_xref="taxon:5691"
                     /clone="238e05"
BASE COUNT       21 a     10 c      8 g     14 t
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 17;  Length 53;
  Best Local Similarity   80.0%;  Pred. No. 2.5e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 AACATCTGTGGGGTTGGTGT 20
           ||||||| || | ||||| |
Db      43 AACATCTATGTGTTTGGTTT 24


RESULT 13
AZ791858
LOCUS       AZ791858                 68 bp    DNA     linear   GSS 16-FEB-2001
DEFINITION  2M0041L21R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC2M0041L21 R, DNA sequence.
ACCESSION   AZ791858
VERSION     AZ791858.1  GI:12935186
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 68)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0041  row: L  column: 21
            Seq primer: CACACAGGAAACAGCTATGACC
            Class: plasmid ends
            High quality sequence stop: 68.
FEATURES             Location/Qualifiers
     source          1. .68
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC2M0041L21"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
BASE COUNT       13 a      4 c     23 g     28 t
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 17;  Length 68;
  Best Local Similarity   80.0%;  Pred. No. 2.8e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 AACATCTGTGGGGTTGGTGT 20
           ||||| |||| |||  ||||
Db      13 AACATGTGTGTGGTATGTGT 32


RESULT 14
AZ377417
LOCUS       AZ377417                 78 bp    DNA     linear   GSS 02-OCT-2000
DEFINITION  1M0131F14R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0131F14 R, DNA sequence.
ACCESSION   AZ377417
VERSION     AZ377417.1  GI:10491117
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 78)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
```

University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0131 row: F column: 14
Seq primer: CACACAGGAAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 78.

FEATURES Location/Qualifiers
source 1. .78
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0131F14"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptored DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptored mouse DNA was annealed to adaptored vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 24 a 14 c 22 g 18 t
ORIGIN

Query Match 68.0%; Score 13.6; DB 17; Length 78;
Best Local Similarity 80.0%; Pred. No. 2.9e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
Qy       1 AACATCTGTGGGGTTGGTGT 20
           ||||   ||||||||| |||
Db      18 AACAAGAGTGGGGTTGCTGT 37
```

RESULT 15
AZ626002
LOCUS AZ626002 85 bp DNA linear GSS 13-DEC-2000
DEFINITION 1M0466D03F Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC1M0466D03 F, DNA sequence.
ACCESSION AZ626002
VERSION AZ626002.1 GI:11748192
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 85)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D.,Weiss,R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss

University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0466 row: D column: 03
Seq primer: CGTTGTAAAACGACGGCCAGT
Class: plasmid ends
High quality sequence stop: 85.

FEATURES Location/Qualifiers
source 1. .85
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0466D03"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptored DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptored mouse DNA was annealed to adaptored vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 18 a 16 c 19 g 32 t
ORIGIN

Query Match 68.0%; Score 13.6; DB 17; Length 85;
Best Local Similarity 80.0%; Pred. No. 3e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
Qy       1 AACATCTGTGGGGTTGGTGT 20
           ||||||||  | ||| ||||
Db      32 AACATCTGGAGTGTTTGTGT 51
```

Search completed: November 26, 2002, 17:57:19
Job time : 809.5 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 3, 2002, 18:14:22 ; Search time 351.3 Seconds
(without alignments)
1656.863 Million cell updates/sec

Title: US-09-296-264-23
Perfect score: 20
Sequence: 1 tcggacaaatcgagttatca 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 995600

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
c	1	14.4	72.0	28	6	AR000028	AR000028 Sequence
c	2	14.4	72.0	28	6	AR169139	AR169139 Sequence
c	3	14.4	72.0	28	6	AR202626	AR202626 Sequence
c	4	14.4	72.0	28	6	AX036240	AX036240 Sequence
	5	14.2	71.0	44	10	RATMLC134	K02426 Rat fast my
	6	13.6	68.0	24	6	E05470	E05470 Synthetic D
	7	13.2	66.0	33	6	AX455866	AX455866 Sequence
	8	12.8	64.0	33	6	AR095659	AR095659 Sequence
	9	12.6	63.0	60	12	SYNBSUA1	M11006 Chimeric pl
	10	12.6	63.0	69	12	SYNBSUB1	M11008 Chimeric pl
	11	12.6	63.0	69	12	SYNBSUC1	M11010 Chimeric pl
	12	12.6	63.0	96	6	AR140870	AR140870 Sequence
	13	12.6	63.0	96	6	AR150820	AR150820 Sequence
	14	12.6	63.0	96	6	I65698	I65698 Sequence 58
	15	12.6	63.0	96	6	I67930	I67930 Sequence 58
	16	12.6	63.0	96	6	I90151	I90151 Sequence 58
	17	12.4	62.0	51	6	AX159849	AX159849 Sequence
	18	12.4	62.0	51	6	AX159850	AX159850 Sequence
	19	12.4	62.0	65	6	AX486221	AX486221 Sequence
	20	12.4	62.0	86	4	AF014842	AF014842 Gulo gulo
c	21	12.2	61.0	17	6	AX423392	AX423392 Sequence
	22	12.2	61.0	18	6	E05471	E05471 Synthetic D
c	23	12.2	61.0	18	6	E05471	E05471 Synthetic D
c	24	12.2	61.0	24	6	AX443743	AX443743 Sequence
c	25	12.2	61.0	24	6	E05470	E05470 Synthetic D
c	26	12.2	61.0	25	6	AX447724	AX447724 Sequence
	27	12.2	61.0	29	6	AR102122	AR102122 Sequence
	28	12.2	61.0	29	6	AR103166	AR103166 Sequence
	29	12.2	61.0	37	6	AX356854	AX356854 Sequence
c	30	12.2	61.0	65	6	AX486112	AX486112 Sequence
	31	12.2	61.0	73	8	MISETRMF	X56005 S.exiguus m
	32	12.2	61.0	90	1	TRNTNA2	M21026 C.perfringe
	33	12.2	61.0	98	8	VFSNU6R	X04788 Broad bean
c	34	12	60.0	42	6	AR154536	AR154536 Sequence
	35	12	60.0	50	6	AR084823	AR084823 Sequence
c	36	12	60.0	50	6	AR084824	AR084824 Sequence
c	37	12	60.0	59	10	AF357318	AF357318 Mus muscu
c	38	12	60.0	65	6	AX483354	AX483354 Sequence
c	39	12	60.0	65	6	AX486279	AX486279 Sequence
	40	12	60.0	65	6	AX486354	AX486354 Sequence
c	41	12	60.0	66	9	HSZ83330	Z83330 H.sapiens s
c	42	12	60.0	68	17	HSMC23D06	X88266 H.sapiens D
	43	12	60.0	74	6	AR102087	AR102087 Sequence
	44	12	60.0	74	6	AR103133	AR103133 Sequence
	45	12	60.0	86	9	HUMDP145	M89843 Human antig

ALIGNMENTS

RESULT 1
AR000028/c
LOCUS AR000028 28 bp DNA linear PAT 04-DEC-1998
DEFINITION Sequence 19 from patent US 5736139.
ACCESSION AR000028
VERSION AR000028.1 GI:3962559
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 28)
AUTHORS Kink,J.A., Thalley,B.S., Stafford,D.C., Firca,J.R. and Padhye,N.V.
TITLE Treatment of Clostridium difficile induced disease
JOURNAL Patent: US 5736139-A 19 07-APR-1998;
FEATURES Location/Qualifiers

```
     source          1. .28
                     /organism="unknown"
BASE COUNT        7 a      4 c      7 g     10 t
ORIGIN

  Query Match            72.0%;  Score 14.4;  DB 6;  Length 28;
  Best Local Similarity   93.8%;  Pred. No. 1.7e+03;
  Matches   15;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       5 ACAAATCGAGTTATCA 20
           ||||||| ||||||||
Db      25 ACAAATCCAGTTATCA 10


RESULT 2
AR169139/c
LOCUS       AR169139                 28 bp    DNA     linear   PAT 17-DEC-2001
DEFINITION  Sequence 19 from patent US 6290960.
ACCESSION   AR169139
VERSION     AR169139.1  GI:17906908
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 28)
  AUTHORS   Kink,J.A., Thalley,B.S. and Stafford,D.C.
  TITLE     Vaccine and antitoxin for the treatment of C. difficile disease
  JOURNAL   Patent: US 6290960-A 19 18-SEP-2001;
FEATURES             Location/Qualifiers
     source          1. .28
                     /organism="unknown"
BASE COUNT        7 a      4 c      7 g     10 t
ORIGIN

  Query Match            72.0%;  Score 14.4;  DB 6;  Length 28;
  Best Local Similarity   93.8%;  Pred. No. 1.7e+03;
  Matches   15;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       5 ACAAATCGAGTTATCA 20
           ||||||| ||||||||
Db      25 ACAAATCCAGTTATCA 10


RESULT 3
AR202626/c
LOCUS       AR202626                 28 bp    DNA     linear   PAT 20-JUN-2002
DEFINITION  Sequence 19 from patent US 6365158.
ACCESSION   AR202626
VERSION     AR202626.1  GI:21498804
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 28)
  AUTHORS   Williams,J.A. and Kink,J.A.
  TITLE     Methods for producing neutralizing antitoxin to C. difficile toxin
            B
  JOURNAL   Patent: US 6365158-A 19 02-APR-2002;
FEATURES             Location/Qualifiers
     source          1. .28
                     /organism="unknown"
BASE COUNT        7 a      4 c      7 g     10 t
ORIGIN

  Query Match            72.0%;  Score 14.4;  DB 6;  Length 28;
  Best Local Similarity   93.8%;  Pred. No. 1.7e+03;
  Matches   15;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       5 ACAAATCGAGTTATCA 20
           ||||||| ||||||||
Db      25 ACAAATCCAGTTATCA 10


RESULT 4
AX036240/c
LOCUS       AX036240                 28 bp    DNA     linear   PAT 16-NOV-2000
DEFINITION  Sequence 19 from Patent EP1041149.
ACCESSION   AX036240
VERSION     AX036240.1  GI:11225861
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 28)
  AUTHORS   Kink,J.A., Firca,J.R., Padhye,N.V., Thalley,B.S., Stafford,D.C. and
            Williams,J.A.
  TITLE     Vaccine and antitoxin for treatment and prevention of c. Difficile
            disease
  JOURNAL   Patent: EP 1041149-A 19 04-OCT-2000;
            OPHIDIAN PHARM INC (US)
FEATURES             Location/Qualifiers
     source          1. .28
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="Synthetic"
BASE COUNT        7 a      4 c      7 g     10 t
ORIGIN

  Query Match            72.0%;  Score 14.4;  DB 6;  Length 28;
  Best Local Similarity   93.8%;  Pred. No. 1.7e+03;
  Matches   15;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       5 ACAAATCGAGTTATCA 20
           ||||||| ||||||||
Db      25 ACAAATCCAGTTATCA 10


RESULT 5
RATMLC134
LOCUS       RATMLC134                44 bp    DNA     linear   ROD 29-APR-1996
DEFINITION  Rat fast myosin alkali light chain exon 2, specific for MLC1-f.
ACCESSION   K02426
VERSION     K02426.1  GI:205468
KEYWORDS    alternative splicing; myosin; myosin light chain.
SEGMENT     4 of 8
SOURCE      Rattus norvegicus (clone:  [pMLC-5,91]and) skeletal muscle cDNA to
            mRNA; and Rattus norvegicus (clone: lambda-[LCH13,14]) (clone
            library: HaeII library) liver DNA.
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 44)
  AUTHORS   Periasamy,M., Strehler,E.E., Garfinkel,L.I., Gubits,R.M.,
            Ruiz-Opazo,N. and Nadal-Ginard,B.
  TITLE     Fast skeletal muscle myosin light chains 1 and 3 are produced from
            a single gene by a combined process of differential RNA
            transcription and splicing
  JOURNAL   J. Biol. Chem. 259 (21), 13595-13604 (1984)
  MEDLINE   85030494
   PUBMED   6092382
COMMENT     See segment 1 for comments.
FEATURES             Location/Qualifiers
     source          1. .44
                     /organism="Rattus norvegicus"
                     /db_xref="taxon:10116"
                     /clone="lambda-[LCH13,14]"
                     /tissue_type="liver"
                     /clone_lib="HaeII library"
     source          11. .39
                     /organism="Rattus norvegicus"
                     /db_xref="taxon:10116"
                     /clone="[pMLC-5,91]and"
                     /tissue_type="skeletal muscle"
     intron          <1. .>44
                     /gene="MLC3-f"
```

```
                  /number=2
     intron       <1. .10
                  /gene="MLC1-f"
                  /number=1
     exon         11. .38
                  /gene="MLC1-f"
                  /number=2
     intron       39. .>44
                  /gene="MLC1-f"
                  /number=2
BASE COUNT       14 a      8 c     12 g     10 t
ORIGIN      About 11 kb after segment 1; about 4.5 kb after segment 3.

  Query Match             71.0%;  Score 14.2;  DB 10;  Length 44;
  Best Local Similarity   84.2%;  Pred. No. 2.4e+03;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 TCGGACAAATCGAGTTATC 19
           || |||| |||||||| ||
Db       3 TCCGACAGATCGAGTTCTC 21


RESULT 6
E05470
LOCUS       E05470                   24 bp    DNA     linear   PAT 29-SEP-1997
DEFINITION  Synthetic DNA sequence.
ACCESSION   E05470
VERSION     E05470.1  GI:2173659
KEYWORDS    JP 1993244960-A/3.
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 24)
  AUTHORS   Murata,T., Horinouchi,S. and Beppu,T.
  TITLE     ASPARTYLPHENYLALANINE POLYMER GENE AND RELATED SUBSTANCE
  JOURNAL   Patent: JP 1993244960-A 3 24-SEP-1993;
            AJINOMOTO CO INC
COMMENT     OS   Artificial gene
            OC   Artificial sequence; Genes.
            PN   JP 1993244960-A/3
            PD   24-SEP-1993
            PF   18-FEB-1992 JP 1992030821
            PI   MURATA TAKUYA, HORINOUCHI SUEJI, BEPPU TERUHIKO PC
            C12N15/62,A23L1/236,C12N1/21,C12N15/71,C12P21/02,(C12N1/21, PC
            C12R1:19),
            PC   (C12P21/02,C12R1:19);
            CC   strandedness: Single;
            CC   topology: Linear;
            CC   hypothetical: No;
            CC   anti-sense: No;
            FH   Key             Location/Qualifiers
            FH
            FT   misc_feature    1. .24
            FT                   /note='synthetic DNA for termination codon and
             FT                   linker'.
FEATURES             Location/Qualifiers
     source          1. .24
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
BASE COUNT        7 a      5 c      4 g      8 t
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 6;  Length 24;
  Best Local Similarity   80.0%;  Pred. No. 5.1e+03;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 TCGGACAAATCGAGTTATCA 20
           || || || |||| ||||||
Db       5 TCTGATAAGTCGACTTATCA 24


RESULT 7
AX455866
LOCUS       AX455866                 33 bp    DNA     linear   PAT 06-JUL-2002
DEFINITION  Sequence 2 from Patent WO0212480.
ACCESSION   AX455866
VERSION     AX455866.1  GI:21714861
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Gao,J., Skeen,R.S., Hooker,B.S. and Anderson,D.B.
  TITLE     Yeast alpha-amylase promoter and uses thereof
  JOURNAL   Patent: WO 0212480-A 2 14-FEB-2002;
            Batelle Memorial Institute (US)
FEATURES             Location/Qualifiers
     source          1. .33
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="oligonucleotide primer"
BASE COUNT       11 a      7 c      5 g     10 t
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 6;  Length 33;
  Best Local Similarity   83.3%;  Pred. No. 9.2e+03;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       3 GGACAAATCGAGTTATCA 20
           | || |||||| ||||||
Db       8 GCACCAATCGATTTATCA 25


RESULT 8
AR095659
LOCUS       AR095659                 33 bp    DNA     linear   PAT 08-SEP-2000
DEFINITION  Sequence 10 from patent US 6004797.
ACCESSION   AR095659
VERSION     AR095659.1  GI:10023732
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 33)
  AUTHORS   Colosi,P.C.
  TITLE     Adenovirus helper-free recombinant AAV Virion production
  JOURNAL   Patent: US 6004797-A 10 21-DEC-1999;
FEATURES             Location/Qualifiers
     source          1. .33
                     /organism="unknown"
BASE COUNT        8 a     10 c      9 g      6 t
ORIGIN

  Query Match             64.0%;  Score 12.8;  DB 6;  Length 33;
  Best Local Similarity   87.5%;  Pred. No. 1.6e+04;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       4 GACAAATCGAGTTATC 19
           |||||||| || ||||
Db       4 GACAAATCTAGATATC 19


RESULT 9
SYNBSUA1
LOCUS       SYNBSUA1                 60 bp    DNA     circular SYN 27-APR-1993
DEFINITION  Chimeric plasmid, mouse interferon-beta (IFN-beta) cDNA fused to
            B.subtilis alpha-amylase signal sequence DNA, 5' junction.
ACCESSION   M11006
VERSION     M11006.1  GI:208059
KEYWORDS    .
SEGMENT     1 of 2
SOURCE      Mouse cDNA to mRNA and B.subtilis alpha-amylase DNA, clone pTUB502.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 60)
  AUTHORS   Shiroza,T., Nakazawa,K., Tashiro,N., Yamane,K., Yanagi,K.,
```

```
          Yamasaki,M., Tamura,G., Saito,H., Kawade,Y. and Taniguchi,T.
  TITLE     Synthesis and secretion of biologically active mouse
            interferon-beta using a Bacillus subtilis alpha-amylase secretion
            vector
  JOURNAL   Gene 34 (1), 1-8 (1985)
  MEDLINE   85232042
   PUBMED   3924734
FEATURES             Location/Qualifiers
     source          1. .60
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
     CDS             <1. .>60
                     /note="alpha-amylase/interferon-beta fusion protein"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="AAA72481.1"
                     /db_xref="GI:208062"
                     /translation="SAETANKSNEQAWLQLQERT"
BASE COUNT       21 a     14 c     17 g      8 t
ORIGIN

  Query Match           63.0%;  Score 12.6;  DB 12;  Length 60;
  Best Local Similarity   78.9%;  Pred. No. 2.3e+04;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      2 CGGACAAATCGAGTTATCA 20
          || ||||||||| | | ||
Db     14 CGAACAAATCGAATGAGCA 32


RESULT 10
SYNBSUB1
LOCUS       SYNBSUB1                 69 bp    DNA     circular SYN 27-APR-1993
DEFINITION  Chimeric plasmid, mouse interferon-beta (IFN-beta) cDNA fused to
            B.subtilis alpha-amylase signal sequence DNA, 5' junction.
ACCESSION   M11008
VERSION     M11008.1  GI:208064
KEYWORDS    .
SEGMENT     1 of 2
SOURCE      Mouse cDNA to mRNA and B.subtilis alpha-amylase DNA, clone pTUB506.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 69)
  AUTHORS   Shiroza,T., Nakazawa,K., Tashiro,N., Yamane,K., Yanagi,K.,
            Yamasaki,M., Tamura,G., Saito,H., Kawade,Y. and Taniguchi,T.
  TITLE     Synthesis and secretion of biologically active mouse
            interferon-beta using a Bacillus subtilis alpha-amylase secretion
            vector
  JOURNAL   Gene 34 (1), 1-8 (1985)
  MEDLINE   85232042
   PUBMED   3924734
FEATURES             Location/Qualifiers
     source          1. .69
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
     CDS             <1. .>69
                     /note="alpha-amylase/interferon-beta fusion protein"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="AAA72483.1"
                     /db_xref="GI:208067"
                     /translation="SAETANKSNEQACYKQLQLQERT"
BASE COUNT       25 a     16 c     18 g     10 t
ORIGIN

  Query Match           63.0%;  Score 12.6;  DB 12;  Length 69;
  Best Local Similarity   78.9%;  Pred. No. 2.3e+04;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      2 CGGACAAATCGAGTTATCA 20
          || ||||||||| | | ||
Db     14 CGAACAAATCGAATGAGCA 32


RESULT 11
SYNBSUC1
LOCUS       SYNBSUC1                 69 bp    DNA     circular SYN 27-APR-1993
DEFINITION  Chimeric plasmid, mouse beta-interferon (beta-IFN) cDNA fused to
            B.subtilis alpha-amylase signal sequence DNA, 5' junction.
ACCESSION   M11010
VERSION     M11010.1  GI:208069
KEYWORDS    .
SEGMENT     1 of 2
SOURCE      Mouse cDNA to mRNA and B.subtilis DNA, clone pTUB509.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 69)
  AUTHORS   Shiroza,T., Nakazawa,K., Tashiro,N., Yamane,K., Yanagi,K.,
            Yamasaki,M., Tamura,G., Saito,H., Kawade,Y. and Taniguchi,T.
  TITLE     Synthesis and secretion of biologically active mouse
            interferon-beta using a Bacillus subtilis alpha-amylase secretion
            vector
  JOURNAL   Gene 34 (1), 1-8 (1985)
  MEDLINE   85232042
   PUBMED   3924734
FEATURES             Location/Qualifiers
     source          1. .69
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
     CDS             <1. .>69
                     /note="alpha-amylase/interferon-beta fusion protein"
                     /codon_start=1
                     /transl_table=11
                     /protein_id="AAA72485.1"
                     /db_xref="GI:208072"
                     /translation="SAETANKSNEHINYKQLQLQERT"
BASE COUNT       27 a     17 c     16 g      9 t
ORIGIN

  Query Match           63.0%;  Score 12.6;  DB 12;  Length 69;
  Best Local Similarity   78.9%;  Pred. No. 2.3e+04;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      2 CGGACAAATCGAGTTATCA 20
          || ||||||||| | | ||
Db     14 CGAACAAATCGAATGAGCA 32


RESULT 12
AR140870
LOCUS       AR140870                 96 bp    DNA     linear   PAT 16-JUN-2001
DEFINITION  Sequence 147 from patent US 6207816.
ACCESSION   AR140870
VERSION     AR140870.1  GI:14483366
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 96)
  AUTHORS   Gold,L., Janjic,N. and Pagratis,N.
  TITLE     High affinity oligonucleotide ligands to growth factors
  JOURNAL   Patent: US 6207816-A 147 27-MAR-2001;
FEATURES             Location/Qualifiers
     source          1. .96
                     /organism="unknown"
BASE COUNT       30 a     24 c     27 g     13 t      2 others
ORIGIN

  Query Match           63.0%;  Score 12.6;  DB 6;  Length 96;
  Best Local Similarity   78.9%;  Pred. No. 2.4e+04;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      2 CGGACAAATCGAGTTATCA 20
          | ||| |||||| | ||||
Db     35 CTGACTAATCGACTGATCA 53
```

```
RESULT 13
AR150820
LOCUS       AR150820                96 bp    DNA     linear   PAT 08-AUG-2001
DEFINITION  Sequence 58 from patent US 6229002.
ACCESSION   AR150820
VERSION     AR150820.1  GI:15115411
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 96)
  AUTHORS   Janjic,N. and Gold,L.
  TITLE     Platelet derived growth factor (PDGF) nucleic acid ligand complexes
  JOURNAL   Patent: US 6229002-A 58 08-MAY-2001;
FEATURES             Location/Qualifiers
     source          1. .96
                     /organism="unknown"
BASE COUNT       30 a     24 c     27 g     13 t      2 others
ORIGIN

  Query Match            63.0%;  Score 12.6;  DB 6;  Length 96;
  Best Local Similarity  78.9%;  Pred. No. 2.4e+04;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 CGGACAAATCGAGTTATCA 20
           | ||| |||||| | ||||
Db      35 CTGACTAATCGACTGATCA 53


RESULT 14
I65698
LOCUS       I65698                  96 bp    DNA     linear   PAT 07-OCT-1997
DEFINITION  Sequence 58 from patent US 5668264.
ACCESSION   I65698
VERSION     I65698.1  GI:2482268
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 96)
  AUTHORS   Janjic,N. and Gold,L.
  TITLE     High affinity PDGF nucleic acid ligands
  JOURNAL   Patent: US 5668264-A 58 16-SEP-1997;
FEATURES             Location/Qualifiers
     source          1. .96
                     /organism="unknown"
BASE COUNT       30 a     24 c     27 g     13 t      2 others
ORIGIN

  Query Match            63.0%;  Score 12.6;  DB 6;  Length 96;
  Best Local Similarity  78.9%;  Pred. No. 2.4e+04;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 CGGACAAATCGAGTTATCA 20
           | ||| |||||| | ||||
Db      35 CTGACTAATCGACTGATCA 53


RESULT 15
I67930
LOCUS       I67930                  96 bp    DNA     linear   PAT 04-FEB-1998
DEFINITION  Sequence 58 from patent US 5674685.
ACCESSION   I67930
VERSION     I67930.1  GI:2830052
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 96)
  AUTHORS   Janjic,N. and Gold,L.
  TITLE     High affinity PDGF nucleic acid ligands
  JOURNAL   Patent: US 5674685-A 58 07-OCT-1997;
FEATURES             Location/Qualifiers
     source          1. .96
                     /organism="unknown"
BASE COUNT       30 a     24 c     27 g     13 t      2 others
ORIGIN

  Query Match            63.0%;  Score 12.6;  DB 6;  Length 96;
  Best Local Similarity  78.9%;  Pred. No. 2.4e+04;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 CGGACAAATCGAGTTATCA 20
           | ||| |||||| | ||||
Db      35 CTGACTAATCGACTGATCA 53


Search completed: December  3, 2002, 22:23:13
Job time : 357.3 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 06:29:45 ; Search time 94.1 Seconds
(without alignments)
478.639 Million cell updates/sec

Title: US-09-296-264-23
Perfect score: 20
Sequence: 1 tcggacaaatcgagttatca 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 2390332

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	20	100.0	20	21	AAZ31453	Human neuropilin m
c	2	14.4	72.0	28	17	AAT29261	C. difficile toxin
c	3	14.4	72.0	28	19	AAV30570	Clostridium diffic
	4	14.2	71.0	52	24	AAL45053	Expression vector
	5	13.8	69.0	65	24	ABN56765	Mouse spliced tran
	6	13.6	68.0	24	14	AAQ50330	Prochymosin fusion
	7	13.2	66.0	24	22	AAH62970	Shrimp white spot
	8	13.2	66.0	33	24	ABK13972	PCR primer AM2-N60
	9	13.2	66.0	60	24	ABN32362	Human spliced tran
	10	13	65.0	28	20	AAZ30784	Oligonucleotide SE
	11	13	65.0	38	20	AAZ30781	Oligonucleotide SE
	12	13	65.0	38	20	AAZ30791	Oligonucleotide SE
	13	13	65.0	39	22	AAD21002	Bovine enterovirus
c	14	12.6	63.0	24	24	ABL54248	Human nucleotide e
c	15	12.6	63.0	29	21	AAF06592	Hammerhead ribozym
c	16	12.6	63.0	52	21	AAZ45712	PCR primer FJ used
	17	12.6	63.0	60	24	ABN47554	Human spliced tran
c	18	12.6	63.0	60	24	ABN48688	Human spliced tran
	19	12.6	63.0	70	20	AAX87206	PCR primer towards
	20	12.6	63.0	70	22	AAD21536	4-4-20 single chai
	21	12.6	63.0	70	24	AAD27161	Single chain varia
c	22	12.6	63.0	71	21	AAC69908	TGF-beta-1-binding
	23	12.6	63.0	96	18	AAT65271	Platelet derived g
	24	12.6	63.0	96	20	AAX87058	Platelet derived g
c	25	12.4	62.0	30	20	AAX88207	Alpha-amylase olig
	26	12.4	62.0	51	22	AAI76236	Human silent SNP c
	27	12.4	62.0	51	22	AAI76237	Human silent SNP c
c	28	12.2	61.0	17	24	ABK19081	Human ERG DNAzyme
	29	12.2	61.0	18	14	AAQ50331	Prochymosin fusion
c	30	12.2	61.0	18	14	AAQ50331	Prochymosin fusion
c	31	12.2	61.0	24	14	AAQ50330	Prochymosin fusion
c	32	12.2	61.0	24	24	ABQ00191	Oligonucleotide ad
c	33	12.2	61.0	24	24	ABQ04348	Oligonucleotide ad
	34	12.2	61.0	24	24	ABQ04389	Oligonucleotide ad
c	35	12.2	61.0	24	24	ABQ10676	Oligonucleotide ad
	36	12.2	61.0	24	24	ABQ10717	Oligonucleotide ad
c	37	12.2	61.0	25	24	ABQ12212	Oligonucleotide ad
	38	12.2	61.0	25	24	ABQ12253	Oligonucleotide ad
	39	12.2	61.0	29	21	AAA72846	Mouse Ice-4 PCR pr
	40	12.2	61.0	37	24	ABK15658	Rice lipoxygenase
c	41	12.2	61.0	60	24	ABN37868	Human spliced tran
c	42	12.2	61.0	60	24	ABN46236	Human spliced tran
c	43	12.2	61.0	60	24	ABN46766	Human spliced tran
	44	12.2	61.0	90	20	AAZ09221	S. aureus enteroto
c	45	12	60.0	27	14	AAQ34860	PCR primer #17 use

ALIGNMENTS

RESULT 1
AAZ31453
ID AAZ31453 standard; DNA; 20 BP.
XX
AC AAZ31453;
XX
DT 07-FEB-2000 (first entry)
XX
DE Human neuropilin mRNA specific antisense oligo GTI3624.
XX
KW Neuropilin; human; growth; metastasis; tumor; neovascularisation;
KW cancer; papilloma; diabetic retinopathy; antisense; ss.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9955855-A2.
XX
PD 04-NOV-1999.
XX
PF 23-APR-1999; 99WO-CA00324.
XX
PR 23-APR-1998; 98US-0082791.
XX
PA (GENE-) GENESENSE TECHNOLOGIES INC.
XX
PI Wright JA, Young AH, Lee YS;
XX
DR WPI; 2000-023357/02.
XX
PT Antisense oligonucleotides that inhibit neuropilin expression, useful
PT for treating cancer -

```
XX
PS   Claim 4; Page 17; 57pp; English.
XX
CC   Sequences AAZ31431-460 represent antisense oligonucleotides which
CC   inhibit human neuropilin expression. The antisense oligonucleotides can
CC   be used to inhibit the growth or metastasis of a mammalian tumor and
CC   inhibit neovascularisation. The oligonucleotides may be used to treat
CC   various forms of cancers or tumors, such as sarcomas, melanomas,
CC   adenomas, carcinomas of solid tissue, hypoxic tumors, squamous cell
CC   carcinomas of the mouth, throat, larynx and lung, genitourinary cancers
CC   such as cervical and bladder cancer, hematopoietic cancers, colon cancer,
CC   breast cancer, pancreatic cancer, renal cancer, brain cancer, skin
CC   cancer, liver cancer, head and neck cancers, and nervous system cancers,
CC   as well as benign lesions such as papillomas. The methods may be used to
CC   treat neovascularisation disorders such as diabetic retinopathy, and
CC   retinopathy of prematurity and age related macular degeneration.
XX
SQ   Sequence 20 BP; 7 A; 4 C; 4 G; 5 T; 0 other;

  Query Match             100.0%;  Score 20;  DB 21;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 0.27;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 TCGGACAAATCGAGTTATCA 20
           ||||||||||||||||||||
Db       1 TCGGACAAATCGAGTTATCA 20


RESULT 2
AAT29261/c
ID   AAT29261 standard; DNA; 28 BP.
XX
AC   AAT29261;
XX
DT   08-JUL-1996  (first entry)
XX
DE   C. difficile toxin B gene PCR primer P14.
XX
KW   Toxin B; cytotoxin; fusion protein; antitoxin; vaccine; immunogen;
KW   Clostridium difficile; diarrhoea; therapy; diagnosis; primer;
KW   polymerase chain reaction; PCR; ss.
XX
OS   Synthetic.
XX
PN   WO9612802-A1.
XX
PD   02-MAY-1996.
XX
PF   23-OCT-1995;   95WO-US13737.
XX
PR   07-JUN-1995;   95US-0480604.
PR   24-OCT-1994;   94US-0329154.
PR   16-MAR-1995;   95US-0405496.
PR   14-APR-1995;   95US-0422711.
XX
PA   (OPHI-) OPHIDIAN PHARM INC.
XX
PI   Firca JR,  Kink JA,  Padhye NV,  Stafford DC,  Thalley BS;
PI   Williams JA;
XX
DR   WPI; 1996-230603/23.
XX
PT   Fusion proteins comprising non-toxin protein and part of toxin -
PT   useful to form anti-toxins against Clostridium botulinum type A, and
PT   C. difficile type toxins, and to treat C. difficile intoxication,
PT   partic. diarrhoea
XX
PS   Example 18; Page 119; 434pp; English.
XX
CC   PCR primers P14 (AAT29261) and P8 (AAT29256) were used to amplify
CC   genomic DNA of Clostridium difficile VPI strain 10463 coding for
CC   toxin B (see also AAR95011) amino acids 1850-2360, a region
CC   including the toxin B repeats (putative ligand binding domain).
CC   Primer P14 adds an EcoRI site immediately flanking the repeats.
CC   The product was expressed in Escherichia coli transformants as a
CC   soluble fusion protein with maltose binding protein as fusion
CC   partner.  Yields of affinity-purified fusion protein were 20 mg/l
CC   of culture fluid, of which approx. 90% was full-length fusion
CC   protein.  Recombinant toxin A proteins can be used for clostridial
CC   antitoxin prodn.
XX
SQ   Sequence 28 BP; 7 A; 4 C; 7 G; 10 T; 0 other;

  Query Match             72.0%;  Score 14.4;  DB 17;  Length 28;
  Best Local Similarity   93.8%;  Pred. No. 3.1e+02;
  Matches   15;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       5 ACAAATCGAGTTATCA 20
             ||||||| ||||||||
Db      25 ACAAATCCAGTTATCA 10


RESULT 3
AAV30570/c
ID   AAV30570 standard; DNA; 28 BP.
XX
AC   AAV30570;
XX
DT   07-DEC-1998  (first entry)
XX
DE   Clostridium difficile toxin B gene PCR primer P14.
XX
KW   Antitoxin; vaccine; cytotoxin; toxin B; intoxication; immunogen;
KW   pseudomembranous enterocolitis; PCR; primer; ss.
XX
OS   Synthetic.
OS   Clostridium difficile.
XX
PN   WO9808540-A1.
XX
PD   05-MAR-1998.
XX
PF   28-AUG-1997;   97WO-US15394.
XX
PR   28-AUG-1996;   96US-0704159.
XX
PA   (OPHI-) OPHIDIAN PHARM INC.
XX
PI   Thalley BS,  Williams JA;
XX
DR   WPI; 1998-230234/20.
XX
PT   Host cell containing recombinant expression vector encoding
PT   Clostridium botulinum type B or E toxin  - useful to treat humans
PT   and other animals at risk of intoxication with clostridial toxin
XX
PS   Example 18; Page 107; 428pp; English.
XX
CC   PCR primer P14 was used in the PCR amplification of fragments of
CC   the toxin B gene (see AAV30561) of Clostridium difficile.  Genomic
CC   DNA of C. difficile VPI strain 10463 was used as template.  Toxin
CC   B gene fragments were cloned into prokaryotic expression vectors
CC   and toxin B polypeptides were subsequently expressed in E. coli
CC   host cells.  It would be advantageous to use simple and inexpensive
CC   prokaryotic expression systems to produce and purify high levels of
CC   recombinant toxin B for immunisation purposes.  The invention also
CC   relates to recombinant toxin polypeptides of Clostridium botulinum
CC   (see AAW68389-400) and their use as immunogens for the production of
CC   vaccines and antitoxins.
XX
SQ   Sequence 28 BP; 7 A; 4 C; 7 G; 10 T; 0 other;

  Query Match             72.0%;  Score 14.4;  DB 19;  Length 28;
  Best Local Similarity   93.8%;  Pred. No. 3.1e+02;
  Matches   15;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;
```

```
Qy       5 ACAAATCGAGTTATCA 20
           ||||||| ||||||||
Db      25 ACAAATCCAGTTATCA 10


RESULT 4
AAL45053
ID   AAL45053 standard; DNA; 52 BP.
XX
AC   AAL45053;
XX
DT   24-MAY-2002  (first entry)
XX
DE   Expression vector related PCR primer #4.
XX
KW   Expression vector; high expression; screening; PCR; primer; ss.
XX
OS   Unidentified.
XX
PN   JP2001352984-A.
XX
PD   25-DEC-2001.
XX
PF   09-JUN-2000; 2000JP-0174318.
XX
PR   09-JUN-2000; 2000JP-0174318.
XX
PA   (TANB ) TT PHARM INC.
XX
DR   WPI; 2002-191856/25.
XX
PT   An expression vector useful for screening a cell highly expressing a
PT   recombinant protein  -
XX
PS   Example 1; Page 8; 16pp; Japanese.
XX
CC   The present invention relates to an expression vector containing an
CC   expression unit containing a first selective marker gene, a cloning site
CC   for inserting a gene encoding an objective protein and an expression unit
CC   containing a second selective marker gene different from the first
CC   selective marker gene in which the cloning site is inserted in close
CC   vicinity between the two expression units. The expression vector is
CC   used for screening a cell strain highly expressing an objective protein.
CC   The present sequence is a PCR primer used in the exemplification of the
CC   invention.
XX
SQ   Sequence 52 BP; 11 A; 7 C; 16 G; 18 T; 0 other;

  Query Match             71.0%;  Score 14.2;  DB 24;  Length 52;
  Best Local Similarity   84.2%;  Pred. No. 4.3e+02;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 CGGACAAATCGAGTTATCA 20
           | ||| ||| |||||||||
Db      13 CTGACTAATTGAGTTATCA 31


RESULT 5
ABN56765
ID   ABN56765 standard; DNA; 65 BP.
XX
AC   ABN56765;
XX
DT   15-JUL-2002  (first entry)
XX
DE   Mouse spliced transcript detection oligonucleotide SEQ ID NO:29513.
XX
KW   Human; mouse; rat; splice transcript; detection; RNA transcript;
KW   splice variant; transcriptome; oligonucleotide library; ss.
XX
OS   Mus musculus.
XX
PN   WO200210449-A2.
XX
PD   07-FEB-2002.
XX
PF   20-JUL-2001; 2001WO-IB01903.
XX
PR   28-JUL-2000; 2000US-221607P.
PR   02-MAY-2001; 2001US-287724P.
XX
PA   (COMP-) COMPUGEN INC.
XX
PI   Shoshan A,  Wasserman A,  Mintz E,  Mintz L,  Faigler S;
XX
DR   WPI; 2002-257383/30.
XX
PT   New oligonucleotide libraries comprising oligonucleotides which
PT   selectively hybridize to mRNAs transcribed from a transcription unit of
PT   a genome, useful for detecting tissue-, pathology-, and
PT   developmental-specific genes  -
XX
PS   Example 1; SEQ ID 29513; 47pp; English.
XX
CC   The present invention describes oligonucleotide libraries for detecting
CC   messenger RNAs that populate a (sub-)transcriptome, where the
CC   (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC   transcription units that populate a genome. The library comprises
CC   several oligonucleotides, each capable of hybridising selectively to a
CC   set of messenger RNAs transcribed from a given transcription unit of
CC   the genome, which encodes one or more messenger RNA splice variants.
CC   The oligonucleotide libraries are useful for detecting mRNAs from a
CC   biological sample, in expression profiling studies, in qualitatively or
CC   quantitatively characterising the corresponding transcriptome, and in
CC   detecting RNA transcripts and splice variants of human or animal
CC   transcriptomes. The libraries may also be used as specialised mini
CC   libraries to detect transcripts of a sub-transcriptome under a
CC   particular biological or pathological state, and so allowing the
CC   detection of tissue- and pathology-specific genes such as those genes
CC   only expressed in specific tissue under a specific pathological
CC   condition; to detect developmental specific genes; and to detect RNA
CC   transcripts and splice variants of a transcriptome of a patient suffering
CC   from a particular disorder. ABN27253 to ABN59589 represent
CC   oligonucleotide sequences from rats, humans and mice, which are used in
CC   the exemplification of the present invention.
CC   N.B. The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 65 BP; 19 A; 15 C; 16 G; 15 T; 0 other;

  Query Match             69.0%;  Score 13.8;  DB 24;  Length 65;
  Best Local Similarity   88.2%;  Pred. No. 7.2e+02;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       4 GACAAATCGAGTTATCA 20
           || |||| |||||||||
Db      25 GAGAAATGGAGTTATCA 41


RESULT 6
AAQ50330
ID   AAQ50330 standard; DNA; 24 BP.
XX
AC   AAQ50330;
XX
DT   10-MAY-1994  (first entry)
XX
DE   Prochymosin fusion gene construction oligomer #3.
XX
KW   Prochymosin; peptide; aspartic acid; phenylalanine; fusion; peptide;
KW   plasmid; transformation; aspartylphenylalanine; DF; ss.
XX
OS   Synthetic.
XX
PN   JP05244960-A.
```

```
XX
PD   24-SEP-1993.
XX
PF   18-FEB-1992;   92JP-0030821.
XX
PR   18-FEB-1992;   92JP-0030821.
XX
PA   (AJIN ) AJINOMOTO KK.
XX
DR   WPI; 1993-338926/43.
XX
PT   Prochymosin and aspartyl-phenylalanine polymer fusion protein and
PT   DNA - for prodn. of protein using microorganism.
XX
PS   Claim 3; Page 7; 11pp; Japanese.
XX
CC   The sequences given in AAQ50328-31 are nucleotide fragments which
CC   were used in th econstruction of a fusion sequence which contains
CC   DNA encoding a fragment of prochymosin linked to a futher DNA fragment
CC   encoding a peptide in which aspartic acid and phenylalanine are
CC   repeated mutually at the C-terminal of the peptide. The fusion
CC   sequence may be included in a plasmid which is used to transform a
CC   microorganism. This allows preparation of aspartylphenylalanine (DF)
CC   efficiently.
XX
SQ   Sequence 24 BP; 7 A; 5 C; 4 G; 8 T; 0 other;

  Query Match             68.0%;  Score 13.6;  DB 14;  Length 24;
  Best Local Similarity   80.0%;  Pred. No. 8.4e+02;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 TCGGACAAATCGAGTTATCA 20
           || || || |||| ||||||
Db       5 TCTGATAAGTCGACTTATCA 24


RESULT 7
AAH62970
ID   AAH62970 standard; DNA; 24 BP.
XX
AC   AAH62970;
XX
DT   11-SEP-2001  (first entry)
XX
DE   Shrimp white spot Bacilliform virus (WSBV) oligonucleotide 131.
XX
KW   Shrimp white spot Bacilliform virus; WSBV; diagnosis; viral infection;
KW   antiviral agent; gene expression; antisense construct; probe; primer;
KW   transgenic viral resistant shrimp; ss.
XX
OS   White spot syndrome virus.
XX
PN   WO200138351-A2.
XX
PD   31-MAY-2001.
XX
PF   08-NOV-2000; 2000WO-US28888.
XX
PR   24-NOV-1999;   99CN-0124717.
XX
PA   (PENY-) PE CORP NY.
PA   (THIR-) THIRD INST OCEANOGRAPHY STATE OCEANI C A.
PA   (SINO-) SINOGENOMAX CO LTD.
XX
PI   Xu X,  Yang F,  He J,  Pham L,  He M,  Ye Y,  Shen Y,  Kodira C;
XX
DR   WPI; 2001-355877/37.
XX
PT   Primary nucleotide sequence of the shrimp white spot Bacilliform virus
PT   (WSBV), useful for producing viral polypeptides that can be used to
PT   screen for agents that are useful for treating WSBV infection -
XX
PS   Disclosure; Figure 3; 626pp; English.
XX
CC   The invention provides the primary nucleotide sequence of the WSBV genome
CC   (AAH62689), predicted transcript sequences (AAH62689-AAH62839) and
CC   encoded proteins (AAG84910-AAG85051) and oligonucleotide sequences
CC   (AAH62840-63160) suitable for use as primers or probes. The nucleic acid
CC   molecules and proteins of the invention are useful for diagnosis and
CC   monitoring viral infection, in screens for antiviral agents and for
CC   monitoring viral gene expression or activity during a treatment regimen.
CC   The nucleic acid molecules are also useful as antisense constructs to
CC   control viral gene expression in infected cells and tissues and to create
CC   transgenic viral resistant shrimp.
XX
SQ   Sequence 24 BP; 8 A; 5 C; 5 G; 6 T; 0 other;

  Query Match             66.0%;  Score 13.2;  DB 22;  Length 24;
  Best Local Similarity   83.3%;  Pred. No. 1.4e+03;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       3 GGACAAATCGAGTTATCA 20
           ||||||||   |||||||
Db       6 GGACAAATAAGGTTATCA 23


RESULT 8
ABK13972
ID   ABK13972 standard; DNA; 33 BP.
XX
AC   ABK13972;
XX
DT   05-JUN-2002  (first entry)
XX
DE   PCR primer AM2-N60 for cloning yeast alpha-amylase promoter.
XX
KW   Yeast; alpha-amylase promoter; AM; Schwanniomyces castellii;
KW   inducible promoter; starch culture medium; glucose culture medium;
KW   regulation of gene expression; PCR; primer; AM2-N60; ss.
XX
OS   Debaryomyces castellii.
XX
PN   WO200212480-A2.
XX
PD   14-FEB-2002.
XX
PF   03-AUG-2001; 2001WO-US24474.
XX
PR   04-AUG-2000; 2000US-0632313.
PR   02-AUG-2001; 2001US-0921942.
XX
PA   (BATT ) BATTELLE MEMORIAL INST.
XX
PI   Gao J,  Skeen RS,  Hooker BS,  Anderson DB;
XX
DR   WPI; 2002-241756/29.
XX
PT   Novel isolated yeast promoter which is native to Schwanniomyces
PT   castellii, located upstream of, and in control of alpha-amylase gene,
PT   is useful for expressing gene of interest in bacterial, yeast, plant
PT   cell species -
XX
PS   Example 1; Page 8; 26pp; English.
XX
CC   The present invention relates to the isolation of a yeast promoter
CC   which is native to Debaryomyces castellii (also known as Schwanniomyces
CC   castellii) and located upstream of, and in control of an alpha-amylase
CC   (AM) gene. The alpha-amylase promoter is useful for expressing and
CC   regulating a gene of interest in bacterial, yeast, mould, and plant cell
CC   species. The alpha-amylase promoter is an inducible promoter, which can
CC   regulate higher gene expression in the presence of starch, or minimal
CC   gene expression in the presence of glucose. The expression yield can be
CC   increased about ten times when induced in a starch culture medium as
CC   compared to a glucose culture medium. The inducibility of the
CC   alpha-amylase promoter provides an opportunity to regulate native or
CC   foreign gene expression in native or heterologous host strains with an
```

```
CC   inexpensive inducing agent i.e. starch. The present sequence represents
CC   a PCR primer used to clone the alpha-amylase promoter.
XX
SQ   Sequence 33 BP; 11 A; 7 C; 5 G; 10 T; 0 other;

  Query Match             66.0%;  Score 13.2;  DB 24;  Length 33;
  Best Local Similarity   83.3%;  Pred. No. 1.4e+03;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       3 GGACAAATCGAGTTATCA 20
           | || |||||| ||||||
Db       8 GCACCAATCGATTTATCA 25


RESULT 9
ABN32362
ID   ABN32362 standard; DNA; 60 BP.
XX
AC   ABN32362;
XX
DT   15-JUL-2002  (first entry)
XX
DE   Human spliced transcript detection oligonucleotide SEQ ID NO:5110.
XX
KW   Human; mouse; rat; splice transcript; detection; RNA transcript;
KW   splice variant; transcriptome; oligonucleotide library; ss.
XX
OS   Homo sapiens.
XX
PN   WO200210449-A2.
XX
PD   07-FEB-2002.
XX
PF   20-JUL-2001; 2001WO-IB01903.
XX
PR   28-JUL-2000; 2000US-221607P.
PR   02-MAY-2001; 2001US-287724P.
XX
PA   (COMP-) COMPUGEN INC.
XX
PI   Shoshan A,  Wasserman A,  Mintz E,  Mintz L,  Faigler S;
XX
DR   WPI; 2002-257383/30.
XX
PT   New oligonucleotide libraries comprising oligonucleotides which
PT   selectively hybridize to mRNAs transcribed from a transcription unit of
PT   a genome, useful for detecting tissue-, pathology-, and
PT   developmental-specific genes  -
XX
PS   Example 1; SEQ ID 5110; 47pp; English.
XX
CC   The present invention describes oligonucleotide libraries for detecting
CC   messenger RNAs that populate a (sub-)transcriptome, where the
CC   (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC   transcription units that populate a genome. The library comprises
CC   several oligonucleotides, each capable of hybridising selectively to a
CC   set of messenger RNAs transcribed from a given transcription unit of
CC   the genome, which encodes one or more messenger RNA splice variants.
CC   The oligonucleotide libraries are useful for detecting mRNAs from a
CC   biological sample, in expression profiling studies, in qualitatively or
CC   quantitatively characterising the corresponding transcriptome, and in
CC   detecting RNA transcripts and splice variants of human or animal
CC   transcriptomes. The libraries may also be used as specialised mini
CC   libraries to detect transcripts of a sub-transcriptome under a
CC   particular biological or pathological state, and so allowing the
CC   detection of tissue- and pathology-specific genes such as those genes
CC   only expressed in specific tissue under a specific pathological
CC   condition; to detect developmental specific genes; and to detect RNA
CC   transcripts and splice variants of a transcriptome of a patient suffering
CC   from a particular disorder. ABN27253 to ABN59589 represent
CC   oligonucleotide sequences from rats, humans and mice, which are used in
CC   the exemplification of the present invention.
CC   N.B. The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 60 BP; 26 A; 9 C; 19 G; 6 T; 0 other;

  Query Match             66.0%;  Score 13.2;  DB 24;  Length 60;
  Best Local Similarity   83.3%;  Pred. No. 1.5e+03;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 TCGGACAAATCGAGTTAT 18
           | ||||||| |||||| |
Db      42 TGGGACAAAACGAGTTGT 59


RESULT 10
AAZ30784
ID   AAZ30784 standard; DNA; 28 BP.
XX
AC   AAZ30784;
XX
DT   05-JAN-2000  (first entry)
XX
DE   Oligonucleotide SEQ ID 4.
XX
KW   Expression; reduction; repression; delaying; modulation; targetting;
KW   eukaryotic; resistance; pathogen; disease;  oncogene;
KW   transcription factor; cellular metabolism; oligonucleotide; ss.
XX
OS   Synthetic.
XX
PN   WO9949029-A1.
XX
PD   30-SEP-1999.
XX
PF   19-MAR-1999;   99WO-AU00195.
XX
PR   20-MAR-1998;   98AU-0002492.
PR   20-MAR-1998;   98AU-0002499.
XX
PA   (AGGE-) AG-GENE AUSTRALIA LTD.
PA   (QUEE-) STATE QUEENSLAND DEPT PRIMARY IND.
XX
PI   Graham MW,  Rice RN;
XX
DR   WPI; 1999-610751/52.
XX
PT   New method for modifying gene expression to confer resistance of
PT   animals and plants to pathogenic viruses  -
XX
PS   Disclosure; Page 153; 160pp; English.
XX
CC   The invention relates to a method of reducing, repressing or delaying
CC   the expression of a target gene in a cell, tissue or organ by
CC   introducing one or more dispersed nucleic acid molecules or foreign
CC   nucleic acid molecules. Reducing, repressing or delaying the expression
CC   of a target gene is effected by introducing one or more dispersed
CC   nucleic acid molecules or foreign nucleic acid molecules. The foreign
CC   nucleic acid molecules contain multiple copies of a nucleotide sequence
CC   of the target gene or region or complementary region. Conditions are
CC   such that there is sufficient translation of the mRNA product of the
CC   target gene to be modified and the transcription of the mRNA product is
CC   not exclusively repressed or reduced. The method is particularly useful
CC   in the modulation of eukaryotic gene expression, in particular human or
CC   animal gene expression, and in the modulation of expression of genes
CC   derived from vertebrate and invertebrate animals. A variety of traits are
CC   selectable including visible traits, disease-resistance traits and
CC   pathogen-resistance traits. The modulatory effect is applicable to
CC   endogenous genes responsible for cellular metabolism or cellular
CC   transformation including oncogenes, transcription factors and other
CC   genes which encode polypeptides involved in cellular metabolism.
CC   There are no prior arts for modulating the level of expression of a
CC   specific gene in a eukaryotic or prokaryotic organism. This invention
CC   provides the means for doing this particularly by repressing, delaying
```

CC or otherwise reducing gene expression. In addition, the methods should
CC provide general means for phenotypic modification without the concomitant
CC gene-targetting approaches.
CC Note: This sequence is given in the sequence listing but is not
CC mentioned elsewhere in the specification.
XX
SQ Sequence 28 BP; 9 A; 7 C; 7 G; 5 T; 0 other;

Query Match 65.0%; Score 13; DB 20; Length 28;
Best Local Similarity 100.0%; Pred. No. 1.8e+03;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 GGACAAATCGAGT 15
||||||||||||
Db 11 GGACAAATCGAGT 23

RESULT 11
AAZ30781
ID AAZ30781 standard; DNA; 38 BP.
XX
AC AAZ30781;
XX
DT 05-JAN-2000 (first entry)
XX
DE Oligonucleotide SEQ ID 1.
XX
KW Expression; reduction; repression; delaying; modulation; targetting;
KW eukaryotic; resistance; pathogen; disease; oncogene;
KW transcription factor; cellular metabolism; oligonucleotide; ss.
XX
OS Synthetic.
XX
PN WO9949029-A1.
XX
PD 30-SEP-1999.
XX
PF 19-MAR-1999; 99WO-AU00195.
XX
PR 20-MAR-1998; 98AU-0002492.
PR 20-MAR-1998; 98AU-0002499.
XX
PA (AGGE-) AG-GENE AUSTRALIA LTD.
PA (QUEE-) STATE QUEENSLAND DEPT PRIMARY IND.
XX
PI Graham MW, Rice RN;
XX
DR WPI; 1999-610751/52.
XX
PT New method for modifying gene expression to confer resistance of
PT animals and plants to pathogenic viruses -
XX
PS Disclosure; Page 152; 160pp; English.
XX
CC The invention relates to a method of reducing, repressing or delaying
CC the expression of a target gene in a cell, tissue or organ by
CC introducing one or more dispersed nucleic acid molecules or foreign
CC nucleic acid molecules. Reducing, repressing or delaying the expression
CC of a target gene is effected by introducing one or more dispersed
CC nucleic acid molecules or foreign nucleic acid molecules. The foreign
CC nucleic acid molecules contain multiple copies of a nucleotide sequence
CC of the target gene or region or complementary region. Conditions are
CC such that there is sufficient translation of the mRNA product of the
CC target gene to be modified and the transcription of the mRNA product is
CC not exclusively repressed or reduced. The method is particularly useful
CC in the modulation of eukaryotic gene expression, in particular human or
CC animal gene expression, and in the modulation of expression of genes
CC derived from vertebrate and invertebrate animals. A variety of traits are
CC selectable including visible traits, disease-resistance traits and
CC pathogen-resistance traits. The modulatory effect is applicable to
CC endogenous genes responsible for cellular metabolism or cellular
CC transformation including oncogenes, transcription factors and other
CC genes which encode polypeptides involved in cellular metabolism.

CC There are no prior arts for modulating the level of expression of a
CC specific gene in a eukaryotic or prokaryotic organism. This invention
CC provides the means for doing this particularly by repressing, delaying
CC or otherwise reducing gene expression. In addition, the methods should
CC provide general means for phenotypic modification without the concomitant
CC gene-targetting approaches.
CC Note: This sequence is given in the sequence listing but is not
CC mentioned elsewhere in the specification.
XX
SQ Sequence 38 BP; 14 A; 9 C; 9 G; 6 T; 0 other;

Query Match 65.0%; Score 13; DB 20; Length 38;
Best Local Similarity 100.0%; Pred. No. 1.9e+03;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 GGACAAATCGAGT 15
||||||||||||
Db 21 GGACAAATCGAGT 33

RESULT 12
AAZ30791
ID AAZ30791 standard; DNA; 38 BP.
XX
AC AAZ30791;
XX
DT 05-JAN-2000 (first entry)
XX
DE Oligonucleotide SEQ ID 11.
XX
KW Expression; reduction; repression; delaying; modulation; targetting;
KW eukaryotic; resistance; pathogen; disease; oncogene;
KW transcription factor; cellular metabolism; oligonucleotide; ss.
XX
OS Synthetic.
XX
PN WO9949029-A1.
XX
PD 30-SEP-1999.
XX
PF 19-MAR-1999; 99WO-AU00195.
XX
PR 20-MAR-1998; 98AU-0002492.
PR 20-MAR-1998; 98AU-0002499.
XX
PA (AGGE-) AG-GENE AUSTRALIA LTD.
PA (QUEE-) STATE QUEENSLAND DEPT PRIMARY IND.
XX
PI Graham MW, Rice RN;
XX
DR WPI; 1999-610751/52.
XX
PT New method for modifying gene expression to confer resistance of
PT animals and plants to pathogenic viruses -
XX
PS Disclosure; Page 156; 160pp; English.
XX
CC The invention relates to a method of reducing, repressing or delaying
CC the expression of a target gene in a cell, tissue or organ by
CC introducing one or more dispersed nucleic acid molecules or foreign
CC nucleic acid molecules. Reducing, repressing or delaying the expression
CC of a target gene is effected by introducing one or more dispersed
CC nucleic acid molecules or foreign nucleic acid molecules. The foreign
CC nucleic acid molecules contain multiple copies of a nucleotide sequence
CC of the target gene or region or complementary region. Conditions are
CC such that there is sufficient translation of the mRNA product of the
CC target gene to be modified and the transcription of the mRNA product is
CC not exclusively repressed or reduced. The method is particularly useful
CC in the modulation of eukaryotic gene expression, in particular human or
CC animal gene expression, and in the modulation of expression of genes
CC derived from vertebrate and invertebrate animals. A variety of traits are
CC selectable including visible traits, disease-resistance traits and
CC pathogen-resistance traits. The modulatory effect is applicable to

CC endogenous genes responsible for cellular metabolism or cellular
CC transformation including oncogenes, transcription factors and other
CC genes which encode polypeptides involved in cellular metabolism.
CC There are no prior arts for modulating the level of expression of a
CC specific gene in a eukaryotic or prokaryotic organism. This invention
CC provides the means for doing this particularly by repressing, delaying
CC or otherwise reducing gene expression. In addition, the methods should
CC provide general means for phenotypic modification without the concomitant
CC gene-targetting approaches.
CC Note: This sequence is given in the sequence listing but is not
CC mentioned elsewhere in the specification.
XX
SQ Sequence 38 BP; 14 A; 9 C; 9 G; 6 T; 0 other;

Query Match 65.0%; Score 13; DB 20; Length 38;
Best Local Similarity 100.0%; Pred. No. 1.9e+03;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 GGACAAATCGAGT 15
||||||||||||||
Db 21 GGACAAATCGAGT 33

RESULT 13
AAD21002
ID AAD21002 standard; DNA; 39 BP.
XX
AC AAD21002;
XX
DT 15-JAN-2002 (first entry)
XX
DE Bovine enterovirus (BEV) RNA polymerase amplifying primer, BEV-1.
XX
KW Genetic silencing; medical industry; animal husbandry; PCR primer;
KW veterinary; gene therapy; Bovine enterovirus; BEV; RNA polymerase; ss.
XX
OS Bovine enterovirus.
XX
PN WO200170949-A1.
XX
PD 27-SEP-2001.
XX
PF 16-MAR-2001; 2001WO-AU00297.
XX
PR 17-MAR-2000; 2000AU-0006363.
PR 24-JAN-2001; 2001AU-0002700.
XX
PA (BENI-) BENITEC AUSTRALIA LTD.
PA (QUEE-) STATE QUEENSLAND DEPT PRIMARY IND.
XX
PI Graham MW, Rice RN, Murphy KM, Reed KC;
XX
DR WPI; 2001-596939/67.
XX
PT Genetic construct for generating transgenic animal cells and in gene
PT therapy of vertebrate animal, comprises nucleotide sequence identical
PT to endogenous target sequence in the genome of vertebrate animal cell
PT -
XX
PS Example 13; Page 69; 176pp; English.
XX
CC The invention relates to a method of inducing, promoting or otherwise
CC facilitating a change in the phenotype of an animal cell or group of
CC animal cells. The modulation of phenotypic expression is conveniently
CC accomplished via genotypic manipulation through such means as reducing
CC translation of transcript to proteinaceous product. The ability to
CC induce, promote or otherwise facilitate the silencing of expressible
CC genetic sequences provides a means for modulating the phenotype in,
CC for example, the medical, veterinary and the animal husbandry
CC industries. Genetic construct comprising nucleotide sequence
CC substantially identical to target endogenous sequence of nucleotides
CC in the genome of a vertebrate animal cell, is useful for altering
CC the phenotype of a vertebrate animal cell, where the phenotype is
CC conferred or otherwise facilitated by the expression of an endogenous
CC gene, and also in the generation of animal cells. It is also useful
CC in gene therapy in a vertebrate animal (e.g. human, primate, livestock
CC animal, laboratory test animal or a murine species, avian species, fish
CC or reptile. The present sequence is a PCR primer used to amplify
CC the complete Bovine enterovirus (BEV) RNA polymerase coding region.
CC BEV RNA polymerase is used in the preparation of genetic constructs.
XX
SQ Sequence 39 BP; 14 A; 10 C; 9 G; 6 T; 0 other;

Query Match 65.0%; Score 13; DB 22; Length 39;
Best Local Similarity 100.0%; Pred. No. 1.9e+03;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 GGACAAATCGAGT 15
||||||||||||||
Db 22 GGACAAATCGAGT 34

RESULT 14
ABL54248/c
ID ABL54248 standard; DNA; 24 BP.
XX
AC ABL54248;
XX
DT 15-JUL-2002 (first entry)
XX
DE Human nucleotide excision repair protein 9.24 PCR primer #1.
XX
KW Nucleotide excision repair protein 9.24; human; endocrine;
KW cytostatic; ophthalmological; tumour; xeroderma pigmentosum;
KW Cockayne syndrome; gene therapy; PCR; primer; ss.
XX
OS Homo sapiens.
XX
PN CN1329052-A.
XX
PD 02-JAN-2002.
XX
PF 19-JUN-2000; 2000CN-0116587.
XX
PR 19-JUN-2000; 2000CN-0116587.
XX
PA (BODE-) BODE GENE DEV CO LTD SHANGHAI.
XX
PI Mao Y, Xie Y;
XX
DR WPI; 2002-352896/39.
XX
PT A novel human nucleotide excision repair protein 9.24 useful for
PT treating, e.g., embryonic developmental deformity, tumor and xeroderma
PT pigmentosum -
XX
PS Example 2; Page 18 (Disclosure); 33pp; Chinese.
XX
CC The present invention relates to novel human nucleotide excision
CC repair protein 9.24 (see ABB76063). The protein and its coding
CC sequence are useful for the treatment of embryonic developmental
CC deformity, tumours, xeroderma pigmentosum and Cockayne syndrome.
CC The present sequence is a PCR primer, which was used in an example
CC from the invention.
XX
SQ Sequence 24 BP; 4 A; 7 C; 7 G; 6 T; 0 other;

Query Match 63.0%; Score 12.6; DB 24; Length 24;
Best Local Similarity 78.9%; Pred. No. 2.9e+03;
Matches 15; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 2 CGGACAAATCGAGTTATCA 20
|||||| || ||| ||||
Db 24 CGGACAGATGGAGGCATCA 6

RESULT 15
AAF06592/c
ID AAF06592 standard; RNA; 29 BP.
XX
AC AAF06592;
XX
DT 16-FEB-2001 (first entry)
XX
DE Hammerhead ribozyme #3389.
XX
KW Ribozyme; erythropoietin; granulocyte colony stimulating factor;
KW interferon alpha; ss.
XX
OS Homo sapiens.
XX
PN WO200061729-A2.
XX
PD 19-OCT-2000.
XX
PF 11-APR-2000; 2000WO-US09721.
XX
PR 12-APR-1999; 99US-0129390.
XX
PA (RIBO-) RIBOZYME PHARM INC.
XX
PI Blatt L, Zwick M, Pavco P, McSwiggen J;
XX
DR WPI; 2000-647423/62.
XX
PT Enzymatic and antisense nucleic acid inhibition of repressor genes,
PT useful for producing e.g. granulocyte colony stimulating factor
PT protein, interferon alpha and erythropoietin -
XX
PS Claim 59; Page 134; 164pp; English.
XX
CC The present invention relates to enzymatic and antisense nucleic acid
CC molecules that act as inhibitors of the expression of repressor genes
CC encoding the TR2 Orphan receptor, EAR3/COUP-TF-1, the GATA
CC transcription factor gene, IRF-2 and/or the CAATT Displacement
CC Protein (CDP). Inhibition of the repressors removes prevents
CC inhibition (and consequently increases expression of) genes involved in
CC the production of erythropoietin, granulocyte colony stimulating factor
CC protein and interferon alpha.
XX
SQ Sequence 29 BP; 10 A; 7 C; 6 G; 5 U; 1 other;

Query Match 63.0%; Score 12.6; DB 21; Length 29;
Best Local Similarity 75.0%; Pred. No. 3e+03;
Matches 15; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```
Qy       1 TCGGACAAATCGAGTTATCA 20
           ||||||| ||| | ||||
Db      29 TCGGACATTTCGNCTCATCA 10
```

Search completed: November 23, 2002, 07:03:44
Job time : 97.1 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 06:36:31 ; Search time 21.3 Seconds
(without alignments)
287.959 Million cell updates/sec

Title: US-09-296-264-23
Perfect score: 20
Sequence: 1 tcggacaaatcgagttatca 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 687286

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	14.4	72.0	28	1	US-08-480-604A-19	Sequence 19, Appl
c	2	14.4	72.0	28	2	US-08-405-496A-19	Sequence 19, Appl
c	3	14.4	72.0	28	4	US-08-915-136-19	Sequence 19, Appl
c	4	14.4	72.0	28	4	US-08-957-310-19	Sequence 19, Appl
	5	12.8	64.0	33	3	US-08-745-957-10	Sequence 10, Appl
	6	12.6	63.0	70	4	US-09-140-084-17	Sequence 17, Appl
	7	12.6	63.0	70	4	US-09-724-297-17	Sequence 17, Appl
c	8	12.6	63.0	71	4	US-09-275-850-103	Sequence 103, App
	9	12.6	63.0	96	1	US-08-479-783A-58	Sequence 58, Appl
	10	12.6	63.0	96	1	US-08-479-725-58	Sequence 58, Appl
	11	12.6	63.0	96	1	US-08-618-693-58	Sequence 58, Appl
	12	12.6	63.0	96	4	US-08-973-124-147	Sequence 147, App
	13	12.6	63.0	96	4	US-08-991-743C-58	Sequence 58, Appl
	14	12.6	63.0	96	5	PCT-US96-08014-147	Sequence 147, App
	15	12.2	61.0	29	3	US-08-258-287B-61	Sequence 61, Appl
	16	12.2	61.0	29	3	US-08-368-704C-59	Sequence 59, Appl
c	17	12	60.0	42	4	US-09-156-828B-16	Sequence 16, Appl
	18	12	60.0	50	2	US-09-060-828A-7	Sequence 7, Appli
c	19	12	60.0	50	2	US-09-060-828A-8	Sequence 8, Appli
	20	12	60.0	74	3	US-08-258-287B-11	Sequence 11, Appl
	21	12	60.0	74	3	US-08-368-704C-11	Sequence 11, Appl
	22	11.6	58.0	25	1	US-08-286-748B-12	Sequence 12, Appl
c	23	11.6	58.0	33	2	US-08-377-309-15	Sequence 15, Appl
c	24	11.6	58.0	33	4	US-09-186-723-15	Sequence 15, Appl
c	25	11.6	58.0	33	4	US-08-505-012-20	Sequence 20, Appl
c	26	11.6	58.0	33	4	US-09-186-949A-16	Sequence 16, Appl
c	27	11.6	58.0	33	5	PCT-US96-00996-20	Sequence 20, Appl
c	28	11.6	58.0	34	4	US-09-386-607-10	Sequence 10, Appl
c	29	11.6	58.0	42	4	US-09-693-146-18	Sequence 18, Appl
c	30	11.4	57.0	22	4	US-09-124-238A-23	Sequence 23, Appl
c	31	11.4	57.0	22	4	US-09-721-975-23	Sequence 23, Appl
	32	11.4	57.0	33	4	US-09-189-653-12	Sequence 12, Appl
	33	11.4	57.0	35	1	US-08-032-846-6	Sequence 6, Appli
	34	11.4	57.0	35	4	US-08-483-511-22	Sequence 22, Appl
	35	11.4	57.0	35	4	US-08-474-636-6	Sequence 6, Appli
	36	11.4	57.0	35	5	PCT-US93-01009-22	Sequence 22, Appl
	37	11.4	57.0	36	1	US-08-741-881-124	Sequence 124, App
	38	11.4	57.0	36	1	US-08-739-158-124	Sequence 124, App
	39	11.4	57.0	36	2	US-08-739-167-124	Sequence 124, App
	40	11.4	57.0	36	3	US-08-404-796-124	Sequence 124, App
	41	11.4	57.0	36	3	US-08-931-869-124	Sequence 124, App
	42	11.4	57.0	36	4	US-09-350-399-124	Sequence 124, App
	43	11.4	57.0	36	4	US-09-236-140A-124	Sequence 124, App
	44	11.4	57.0	39	3	US-08-776-900C-21	Sequence 21, Appl
	45	11.4	57.0	39	4	US-09-268-195C-21	Sequence 21, Appl

ALIGNMENTS

```
RESULT 1
US-08-480-604A-19/c
; Sequence 19, Application US/08480604A
; Patent No. 5736139
;  GENERAL INFORMATION:
;    APPLICANT:  KINK, JOHN A.
;    APPLICANT:  THALLEY, BRUCE S.
;    APPLICANT:  PADHYE, NISHA V.
;    APPLICANT:  FIRCA, JOSEPH R.
;    APPLICANT:  STAFFORD, DOUGLAS C.
;    TITLE OF INVENTION:  VACCINE AND ANTITOXIN FOR TREATMENT AND
;    TITLE OF INVENTION:  PREVENTION OF C. DIFFICILE DISEASE
;    NUMBER OF SEQUENCES:  32
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MEDLEN & CARROLL, LLP
;      STREET:  220 MONTGOMERY STREET, SUITE 2200
;      CITY:  SAN FRANCISCO
;      STATE:  CALIFORNIA
;      COUNTRY:  UNITED STATES OF AMERICA
;      ZIP:  94104
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/480,604A
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/422,711
;      FILING DATE:  14-APR-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/405,496
;      FILING DATE:  16-MAR-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/329,154
;      FILING DATE:  25-OCT-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/161,907
;      FILING DATE:  02-DEC-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/985,321
;      FILING DATE:  04-DEC-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/429,791
;      FILING DATE:  31-OCT-1989
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  INGOLIA, DIANE E.
;      REGISTRATION NUMBER:  40,027
```

```
;      REFERENCE/DOCKET NUMBER:  OPHD-01763
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 705-8410
;      TELEFAX:  (415) 397-8338
;  INFORMATION FOR SEQ ID NO:  19:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  28 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-480-604A-19

  Query Match             72.0%;  Score 14.4;  DB 1;  Length 28;
  Best Local Similarity   93.8%;  Pred. No. 39;
  Matches   15;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       5 ACAAATCGAGTTATCA 20
           ||||||| ||||||||
Db      25 ACAAATCCAGTTATCA 10


RESULT 2
US-08-405-496A-19/c
; Sequence 19, Application US/08405496A
; Patent No. 5919665
;  GENERAL INFORMATION:
;    APPLICANT:  WILLIAMS, JAMES A.
;    TITLE OF INVENTION:  VACCINE FOR CLOSTRIDIUM BOTULINUM
;    TITLE OF INVENTION:  NEUROTOXIN
;    NUMBER OF SEQUENCES:  30
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MEDLEN & CARROLL, LLP
;      STREET:  220 MONTGOMERY STREET, SUITE 2200
;      CITY:  SAN FRANCISCO
;      STATE:  CALIFORNIA
;      COUNTRY:  USA
;      ZIP:  94104
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/405,496A
;      FILING DATE:  16-MAR-1995
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/329,154
;      FILING DATE:  25-OCT-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/161,907
;      FILING DATE:  02-DEC-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/985,321
;      FILING DATE:  04-DEC-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/429,791
;      FILING DATE:  31-OCT-1989
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  INGOLIA, DIANE E.
;      REGISTRATION NUMBER:  40,027
;      REFERENCE/DOCKET NUMBER:  OPHD-01308
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 705-8410
;      TELEFAX:  (415) 397-8338
;  INFORMATION FOR SEQ ID NO:  19:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  28 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-405-496A-19

  Query Match             72.0%;  Score 14.4;  DB 2;  Length 28;
  Best Local Similarity   93.8%;  Pred. No. 39;
  Matches   15;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       5 ACAAATCGAGTTATCA 20
           ||||||| ||||||||
Db      25 ACAAATCCAGTTATCA 10


RESULT 3
US-08-915-136-19/c
; Sequence 19, Application US/08915136
; Patent No. 6290960
;  GENERAL INFORMATION:
;    APPLICANT:  KINK, JOHN A.
;    APPLICANT:  THALLEY, BRUCE S.
;    APPLICANT:  PADHYE, NISHA V.
;    APPLICANT:  FIRCA, JOSEPH R.
;    APPLICANT:  STAFFORD, DOUGLAS C.
;    TITLE OF INVENTION:  VACCINE AND ANTITOXIN FOR TREATMENT AND
;    TITLE OF INVENTION:  PREVENTION OF C. DIFFICILE DISEASE
;    NUMBER OF SEQUENCES:  32
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MEDLEN & CARROLL, LLP
;      STREET:  220 MONTGOMERY STREET, SUITE 2200
;      CITY:  SAN FRANCISCO
;      STATE:  CALIFORNIA
;      COUNTRY:  UNITED STATES OF AMERICA
;      ZIP:  94104
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/915,136
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/480,604
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/405,496
;      FILING DATE:  16-MAR-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/329,154
;      FILING DATE:  25-OCT-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/161,907
;      FILING DATE:  02-DEC-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/985,321
;      FILING DATE:  04-DEC-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/429,791
;      FILING DATE:  31-OCT-1989
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  INGOLIA, DIANE E.
;      REGISTRATION NUMBER:  40,027
;      REFERENCE/DOCKET NUMBER:  OPHD-01763
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 705-8410
;      TELEFAX:  (415) 397-8338
;  INFORMATION FOR SEQ ID NO:  19:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  28 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-915-136-19
```

```
  Query Match             72.0%;  Score 14.4;  DB 4;  Length 28;
  Best Local Similarity   93.8%;  Pred. No. 39;
  Matches   15;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       5 ACAAATCGAGTTATCA 20
           ||||||| ||||||||
Db      25 ACAAATCCAGTTATCA 10


RESULT 4
US-08-957-310-19/c
; Sequence 19, Application US/08957310
; Patent No. 6365158
;  GENERAL INFORMATION:
;    APPLICANT:  Williams, James A.
;    APPLICANT:  Kink, John A.
;    TITLE OF INVENTION:  IDENTIFICATION OF NEUTRALIZING EPITOPES
;    TITLE OF INVENTION:  OF TOXIN A AND TOXIN B FOR THE TREATMENT OF C. DIFFICILE
;    TITLE OF INVENTION:  DISEASE
;    NUMBER OF SEQUENCES:  22
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Medlen & Carroll
;      STREET:  220 Montgomery Street, Suite 2200
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  United States of America
;      ZIP:  94104
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/957,310
;      FILING DATE:  23-OCT-1997
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/329,154
;      FILING DATE:  24-OCT-1994
;      APPLICATION NUMBER:  US 08/161,907
;      FILING DATE:  02-DEC-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/985,321
;      FILING DATE:  04-DEC-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/429,791
;      FILING DATE:  31-OCT-1989
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Ingolia, Diane E.
;      REGISTRATION NUMBER:  40,027
;      REFERENCE/DOCKET NUMBER:  OPHD-01121
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 705-8410
;      TELEFAX:  (415) 397-8338
;  INFORMATION FOR SEQ ID NO:  19:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  28 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-957-310-19

  Query Match             72.0%;  Score 14.4;  DB 4;  Length 28;
  Best Local Similarity   93.8%;  Pred. No. 39;
  Matches   15;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       5 ACAAATCGAGTTATCA 20
           ||||||| ||||||||
Db      25 ACAAATCCAGTTATCA 10


RESULT 5
US-08-745-957-10
; Sequence 10, Application US/08745957
; Patent No. 6004797
;  GENERAL INFORMATION:
;    APPLICANT:  COLOSI, PETER C.
;    TITLE OF INVENTION:  ACCESSORY FUNCTIONS FOR USE IN
;    TITLE OF INVENTION:  RECOMBINANT AAV VIRION PRODUCTION
;    NUMBER OF SEQUENCES:  17
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  REED & ROBINS LLP
;      STREET:  285 HAMILTON AVENUE
;      CITY:  PALO ALTO
;      STATE:  CALIFORNIA
;      COUNTRY:  USA
;      ZIP:  94301
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/745,957
;      FILING DATE:  07-NOV-1996
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 60/006,402
;      FILING DATE:  09-NOV-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  MCCRACKEN, THOMAS P.
;      REGISTRATION NUMBER:  38,548
;      REFERENCE/DOCKET NUMBER:  0800-0007
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 327-3400
;      TELEFAX:  (415)327-3231
;  INFORMATION FOR SEQ ID NO:  10:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  33 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-745-957-10

  Query Match             64.0%;  Score 12.8;  DB 3;  Length 33;
  Best Local Similarity   87.5%;  Pred. No. 3e+02;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       4 GACAAATCGAGTTATC 19
           |||||||| || ||||
Db       4 GACAAATCTAGATATC 19


RESULT 6
US-09-140-084-17
; Sequence 17, Application US/09140084A
; Patent No. 6300065
; GENERAL INFORMATION:
;  APPLICANT: Kieke, et al.
;  TITLE OF INVENTION: Yeast Cell Surface Display of Proteins and Uses Thereof
;  FILE REFERENCE: D6061CIP2
;  CURRENT APPLICATION NUMBER: US/09/140,084A
;  CURRENT FILING DATE: 1998-08-26
;  NUMBER OF SEQ ID NOS: 26
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 17
;   LENGTH: 70
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:PCR Primer
;   OTHER INFORMATION: towards 4-4-20 scFv
US-09-140-084-17
```

```
  Query Match            63.0%;  Score 12.6;  DB 4;  Length 70;
  Best Local Similarity  78.9%;  Pred. No. 4.1e+02;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      2 CGGACAAATCGAGTTATCA 20
          ||| ||| ||||| ||| |
Db      4 CGGCCAACTCGAGCTATTA 22


RESULT 7
US-09-724-297-17
; Sequence 17, Application US/09724297
; Patent No. 6423538
; GENERAL INFORMATION:
;  APPLICANT:  The Board of Trustees of the University of Illinois
;  APPLICANT:  Wittrup, et al.
;  TITLE OF INVENTION:  Yeast Cell Surface Display of Proteins and Uses Thereof
;  FILE REFERENCE:  97-99C
;  CURRENT APPLICATION NUMBER: US/09/724,297
;  CURRENT FILING DATE:  2000-11-28
;  PRIOR APPLICATION NUMBER:  US 09/009,388
;  PRIOR FILING DATE:  1998-01-20
;  NUMBER OF SEQ ID NOS:  26
;  SOFTWARE:  PatentIn version 3.0
; SEQ ID NO 17
;   LENGTH: 70
;   TYPE: DNA
;   ORGANISM: artificial sequence
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: ()..()
;   OTHER INFORMATION: PCR primer towards 4-4-20 scFv
US-09-724-297-17

  Query Match            63.0%;  Score 12.6;  DB 4;  Length 70;
  Best Local Similarity  78.9%;  Pred. No. 4.1e+02;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      2 CGGACAAATCGAGTTATCA 20
          ||| ||| ||||| ||| |
Db      4 CGGCCAACTCGAGCTATTA 22


RESULT 8
US-09-275-850-103/c
; Sequence 103, Application US/09275850A
; Patent No. 6261774
; GENERAL INFORMATION:
;  APPLICANT:  Pagratis, Nikos
;  APPLICANT:  Gold, Larry
;  APPLICANT:  Shtatland, Timur
;  APPLICANT:  Javornik, Brenda
;  TITLE OF INVENTION:  Truncation SELEX Method
;  FILE REFERENCE:  NEX 79
;  CURRENT APPLICATION NUMBER: US/09/275,850A
;  CURRENT FILING DATE:  1999-03-24
;  NUMBER OF SEQ ID NOS:  351
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 103
;   LENGTH: 71
;   TYPE: RNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Synthetic
;   OTHER INFORMATION: Sequence
;   NAME/KEY: modified_base
;   LOCATION: (1)..(71)
;   OTHER INFORMATION: All pyrimidines are 2'-F
US-09-275-850-103

  Query Match            63.0%;  Score 12.6;  DB 4;  Length 71;
  Best Local Similarity  78.9%;  Pred. No. 4.1e+02;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      2 CGGACAAATCGAGTTATCA 20
          ||| ||| ||||| ||| |
Db     55 CGGACAGATCAAGGAATCA 37


RESULT 9
US-08-479-783A-58
; Sequence 58, Application US/08479783A
; Patent No. 5668264
;  GENERAL INFORMATION:
;    APPLICANT:  NEBOJSA JANJIC
;    APPLICANT:  LARRY GOLD
;    TITLE OF INVENTION:  HIGH AFFINITY PDGF NUCLEIC
;    TITLE OF INVENTION:  ACID LIGANDS
;    NUMBER OF SEQUENCES:  90
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Swanson and Bratschun, L.L.C.
;      STREET:  8400 East Prentice Avenue, Suite #200
;      CITY:  Denver
;      STATE:  Colorado
;      COUNTRY:  USA
;      ZIP:  80111
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.5 inch, 1.4 Mb storage
;      COMPUTER:  IBM compatible
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  WordPerfect 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/479,783A
;      FILING DATE:  7-JUNE-1995
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/714,131
;      FILING DATE:  10-JUNE-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/931,473
;      FILING DATE:  17-AUGUST-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/964,624
;      FILING DATE:  21-OCTOBER-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/117,991
;      FILING DATE:  8-SEPTEMBER-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/536,428
;      FILING DATE:  11-JUNE-1990
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Diane H. McClearn
;      REGISTRATION NUMBER:  33,960
;      REFERENCE/DOCKET NUMBER:  NEX42-2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (303) 793-3333
;      TELEFAX:  (303) 793-3433
;  INFORMATION FOR SEQ ID NO:  58:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  96 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  RNA
;    FEATURE:
;      OTHER INFORMATION:  All pyrimidines are 2'-F
;      OTHER INFORMATION:  modified
US-08-479-783A-58

  Query Match            63.0%;  Score 12.6;  DB 1;  Length 96;
  Best Local Similarity  63.2%;  Pred. No. 4.3e+02;
  Matches   12;  Conservative    3;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      2 CGGACAAATCGAGTTATCA 20
          | ||| ||:||| : |:||
```

```
Db      35 CUGACUAAUCGACUGAUCA 53


RESULT 10
US-08-479-725-58
; Sequence 58, Application US/08479725
; Patent No. 5674685
;  GENERAL INFORMATION:
;    APPLICANT:  NEBOJSA JANJIC
;    APPLICANT:  LARRY GOLD
;    TITLE OF INVENTION:  HIGH AFFINITY PDGF NUCLEIC
;    TITLE OF INVENTION:  ACID LIGANDS
;    NUMBER OF SEQUENCES:  90
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Swanson and Bratschun, L.L.C.
;      STREET:  8400 East Prentice Avenue, Suite #200
;      CITY:  Denver
;      STATE:  Colorado
;      COUNTRY:  USA
;      ZIP:  80111
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.5 inch, 1.4 Mb storage
;      COMPUTER:  IBM compatible
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  WordPerfect 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/479,725
;      FILING DATE:  7-JUNE-1995
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/714,131
;      FILING DATE:  10-JUNE-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/931,473
;      FILING DATE:  17-AUGUST-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/964,624
;      FILING DATE:  21-OCTOBER-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/117,991
;      FILING DATE:  8-SEPTEMBER-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/536,428
;      FILING DATE:  11-JUNE-1990
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Diane H. McClearn
;      REGISTRATION NUMBER:  33,960
;      REFERENCE/DOCKET NUMBER:  NEX42-1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (303) 793-3333
;      TELEFAX:  (303) 793-3433
;  INFORMATION FOR SEQ ID NO:  58:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  96 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  RNA
;    FEATURE:
;      OTHER INFORMATION:  All pyrimidines are 2'-F
;      OTHER INFORMATION:  modified
US-08-479-725-58

  Query Match             63.0%;  Score 12.6;  DB 1;  Length 96;
  Best Local Similarity   63.2%;  Pred. No. 4.3e+02;
  Matches   12;  Conservative    3;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 CGGACAAATCGAGTTATCA 20
           | ||| ||:||| : |:||
Db      35 CUGACUAAUCGACUGAUCA 53


RESULT 11
US-08-618-693-58
; Sequence 58, Application US/08618693
; Patent No. 5723594
;  GENERAL INFORMATION:
;    APPLICANT:  NEBOJSA JANJIC
;    APPLICANT:  LARRY GOLD
;    TITLE OF INVENTION:  HIGH AFFINITY PDGF NUCLEIC
;    TITLE OF INVENTION:  ACID LIGANDS
;    NUMBER OF SEQUENCES:  96
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Swanson and Bratschun, L.L.C.
;      STREET:  8400 East Prentice Avenue, Suite #200
;      CITY:  Denver
;      STATE:  Colorado
;      COUNTRY:  USA
;      ZIP:  80111
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.5 inch, 1.4 Mb storage
;      COMPUTER:  IBM compatible
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  WordPerfect 6.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/618,693
;      FILING DATE:  20 MARCH 1996
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/479,783
;      FILING DATE:  7-JUNE-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Barry J. Swanson
;      REGISTRATION NUMBER:  33,215
;      REFERENCE/DOCKET NUMBER:  NEX42/CIP
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (303) 793-3333
;      TELEFAX:  (303) 793-3433
;  INFORMATION FOR SEQ ID NO:  58:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  96 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  RNA
;    FEATURE:
;      OTHER INFORMATION:  All pyrimidines are 2'-F
;      OTHER INFORMATION:  modified
US-08-618-693-58

  Query Match             63.0%;  Score 12.6;  DB 1;  Length 96;
  Best Local Similarity   63.2%;  Pred. No. 4.3e+02;
  Matches   12;  Conservative    3;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 CGGACAAATCGAGTTATCA 20
           | ||| ||:||| : |:||
Db      35 CUGACUAAUCGACUGAUCA 53


RESULT 12
US-08-973-124-147
; Sequence 147, Application US/08973124
; Patent No. 6207816
;  GENERAL INFORMATION:
;    APPLICANT:  LARRY GOLD et al.
;    TITLE OF INVENTION:  HIGH AFFINITY OLIGONUCLEOTIDE
;    TITLE OF INVENTION:  LIGANDS TO GROWTH
;    TITLE OF INVENTION:  FACTORS
;    NUMBER OF SEQUENCES:  304
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Swanson & Bratschun, L.L.C.
;      STREET:  8400 E. Prentice Avenue, Suite 200
;      CITY:  Englewood
;      STATE:  Colorado
;      COUNTRY:  USA
;      ZIP:  80111
```

```
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.5 inch, 1.44 Mb storage
;      COMPUTER:  IBM compatible
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  WordPerfect 6.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/973,124
;      FILING DATE:
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US96/08014
;      FILING DATE:  30-MAY-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/458,423
;      FILING DATE:  02-JUNE-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/458,424
;      FILING DATE:  02-JUNE-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/465,594
;      FILING DATE:  05-JUNE-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/465,591
;      FILING DATE:  05-JUNE-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/479,725
;      FILING DATE:  07-JUNE-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/479,783
;      FILING DATE:  07-JUNE-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/618,693
;      FILING DATE:  20-MARCH-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Barry J. Swanson
;      REGISTRATION NUMBER:  33,215
;      REFERENCE/DOCKET NUMBER:
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (303) 793-3333
;      TELEFAX:  (303) 793-3433
;  INFORMATION FOR SEQ ID NO:  147:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  96 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  RNA
;    FEATURE:
;      OTHER INFORMATION:  All pyrimidines are 2'-F modified
US-08-973-124-147

  Query Match             63.0%;  Score 12.6;  DB 4;  Length 96;
  Best Local Similarity   63.2%;  Pred. No. 4.3e+02;
  Matches   12;  Conservative    3;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 CGGACAAATCGAGTTATCA 20
           | ||| ||:||| : |:||
Db      35 CUGACUAAUCGACUGAUCA 53


RESULT 13
US-08-991-743C-58
; Sequence 58, Application US/08991743C
; Patent No. 6229002
;   GENERAL INFORMATION:
;        APPLICANT: NEBOJSA JANJIC, LARRY GOLD
;        TITLE OF INVENTION: PLATELET DERIVED GROWTH FACTOR (PDGF) NUCLEIC
;                            ACID LIGAND COMPLEXES
;        NUMBER OF SEQUENCES: 149
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Swanson and Bratschun, L.L.C.
;             STREET: 1745 Shea Center Drive, Suite 330
;             CITY: Highlands Ranch
;             STATE: Colorado
;             COUNTRY: USA
;             ZIP: 80129
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Diskette, 3.5 inch, 1.4 Mb storage
;             COMPUTER: IBM compatible
;             OPERATING SYSTEM: MS-DOS
;             SOFTWARE: Word 97
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/08/991,743C
;             FILING DATE: 16-Dec-1997
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 08/618,693
;             FILING DATE: 20-MARCH-1996
;             APPLICATION NUMBER: 08/479,783
;             FILING DATE: 7-JUNE-1995
;             APPLICATION NUMBER: 08/479,725
;             FILING DATE: 7-JUNE-1995
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Barry J. Swanson
;             REGISTRATION NUMBER: 33,215
;             REFERENCE/DOCKET NUMBER: NEX66
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (303) 268-0066
;             TELEFAX: (303) 268-0065
;   INFORMATION FOR SEQ ID NO: 58:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 96 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: RNA
;        FEATURE:
;             OTHER INFORMATION:   All pyrimidines are 2'-F modified
;        SEQUENCE DESCRIPTION: SEQ ID NO: 58:
US-08-991-743C-58

  Query Match             63.0%;  Score 12.6;  DB 4;  Length 96;
  Best Local Similarity   63.2%;  Pred. No. 4.3e+02;
  Matches   12;  Conservative    3;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 CGGACAAATCGAGTTATCA 20
           | ||| ||:||| : |:||
Db      35 CUGACUAAUCGACUGAUCA 53


RESULT 14
PCT-US96-08014-147
; Sequence 147, Application PC/TUS9608014
;  GENERAL INFORMATION:
;    APPLICANT:  LARRY GOLD; NEBOJSA JANJIC; STEVEN RINGQUIST;    NIKOS
;    APPLICANT:  PAGRATIS;PENELOPE J. TOOTHMAN
;    TITLE OF INVENTION:  HIGH AFFINITY OLIGONUCLEOTIDE
;    TITLE OF INVENTION:  LIGANDS TO TRANSFORMING GROWTH
;    TITLE OF INVENTION:  FACTOR   (TGF ), PLATELET-DERIVED
;    TITLE OF INVENTION:  GROWTH FACTOR (PDGF) AND HUMAN
;    TITLE OF INVENTION:  KEROTINOCYTE GROWTH FACTOR (hKGF)
;    NUMBER OF SEQUENCES:  304
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Swanson & Bratschun, L.L.C.
;      STREET:  8400 E. Prentice Avenue, Suite 200
;      CITY:  Englewood
;      STATE:  Colorado
;      COUNTRY:  USA
;      ZIP:  80111
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.5 inch, 1.44 Mb storage
;      COMPUTER:  IBM compatible
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  WordPerfect 6.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US96/08014
;      FILING DATE:
```

```
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/458,423
;      FILING DATE:  02-JUNE-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/458,424
;      FILING DATE:  02-JUNE-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/465,594
;      FILING DATE:  05-JUNE-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/465,591
;      FILING DATE:  05-JUNE-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/479,725
;      FILING DATE:  07-JUNE-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/479,783
;      FILING DATE:  07-JUNE-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/618,693
;      FILING DATE:  20-MARCH-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Barry J. Swanson
;      REGISTRATION NUMBER:  33,215
;      REFERENCE/DOCKET NUMBER:
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (303) 793-3333
;      TELEFAX:  (303) 793-3433
;  INFORMATION FOR SEQ ID NO:  147:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  96 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  RNA
;    FEATURE:
;      OTHER INFORMATION:  All pyrimidines are 2'-F modified
PCT-US96-08014-147

  Query Match             63.0%;  Score 12.6;  DB 5;  Length 96;
  Best Local Similarity   63.2%;  Pred. No. 4.3e+02;
  Matches   12;  Conservative    3;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 CGGACAAATCGAGTTATCA 20
           | ||| ||:||| : |:||
Db      35 CUGACUAAUCGACUGAUCA 53


RESULT 15
US-08-258-287B-61
; Sequence 61, Application US/08258287B
; Patent No. 6083735
;  GENERAL INFORMATION:
;    APPLICANT:  Yuan, Junying
;    APPLICANT:  Miura, Masayuki
;    TITLE OF INVENTION:  Programmed Cell Death Genes and Proteins
;    NUMBER OF SEQUENCES:  85
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sterne, Kessler, Goldstein & Fox
;      STREET:  1100 New York Avenue, Suite 600
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20005
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/258,287B
;      FILING DATE:  10-JUN-1994
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/080,850
;      FILING DATE:  24-JUN-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Bugaisky, Lawrence B.
;      REGISTRATION NUMBER:  35,086
;      REFERENCE/DOCKET NUMBER:  0609.3920001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 371-2600
;      TELEFAX:  (202) 371-2540
;      TELEX:  248636 SSK
;  INFORMATION FOR SEQ ID NO:  61:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  29 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  both
;      TOPOLOGY:  both
US-08-258-287B-61

  Query Match             61.0%;  Score 12.2;  DB 3;  Length 29;
  Best Local Similarity   82.4%;  Pred. No. 6.2e+02;
  Matches   14;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 TCGGACAAATCGAGTTA 17
           ||||  |||||||| ||
Db      10 TCGGGGAAATCGAGGTA 26


Search completed: November 23, 2002, 07:07:34
Job time : 23.3 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 06:42:25 ; Search time 16.8 Seconds
(without alignments)
450.869 Million cell updates/sec

Title: US-09-296-264-23
Perfect score: 20
Sequence: 1 tcggacaaatcgagttatca 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 335578 seqs, 189365133 residues

Total number of hits satisfying chosen parameters: 195614

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	13.2	66.0	33	10	US-09-921-942B-2	Sequence 2, Appli
2	13	65.0	28	9	US-09-100-812-4	Sequence 4, Appli
3	13	65.0	38	9	US-09-100-812-1	Sequence 1, Appli
4	12.2	61.0	96	10	US-09-923-876-827	Sequence 827, App
5	11.6	58.0	31	10	US-09-801-274-833	Sequence 833, App
c 6	11.6	58.0	33	9	US-10-115-701A-15	Sequence 15, Appl
7	11.6	58.0	91	10	US-09-294-093B-1011	Sequence 1011, Ap
8	11.4	57.0	35	9	US-09-466-035-22	Sequence 22, Appl
9	11.4	57.0	35	10	US-09-912-679-22	Sequence 22, Appl
10	11.2	56.0	44	9	US-09-978-295A-181	Sequence 181, App
11	11.2	56.0	44	9	US-09-978-697-181	Sequence 181, App
12	11.2	56.0	77	10	US-09-923-876-3345	Sequence 3345, Ap
13	11.2	56.0	97	10	US-09-878-574-11180	Sequence 11180, A
14	11.2	56.0	99	10	US-09-864-761-32354	Sequence 32354, A
15	11	55.0	35	10	US-09-971-309-85	Sequence 85, Appl
16	11	55.0	36	10	US-09-833-555-29	Sequence 29, Appl
17	11	55.0	39	9	US-09-832-659-21	Sequence 21, Appl
c 18	11	55.0	66	10	US-09-766-113-16	Sequence 16, Appl
c 19	11	55.0	84	10	US-09-864-761-19171	Sequence 19171, A
c 20	11	55.0	86	9	US-09-764-884-3	Sequence 3, Appli
c 21	11	55.0	86	9	US-09-764-857-3	Sequence 3, Appli
c 22	11	55.0	86	9	US-09-860-670-3	Sequence 3, Appli
c 23	11	55.0	86	9	US-09-981-876-3	Sequence 3, Appli
c 24	11	55.0	86	9	US-09-984-245-3	Sequence 3, Appli
c 25	11	55.0	86	9	US-09-764-868-3	Sequence 3, Appli
c 26	11	55.0	86	10	US-09-126-945B-6	Sequence 6, Appli
c 27	11	55.0	86	10	US-09-739-907-3	Sequence 3, Appli
c 28	11	55.0	86	10	US-09-729-835-3	Sequence 3, Appli
c 29	11	55.0	86	10	US-09-739-254-3	Sequence 3, Appli
c 30	11	55.0	86	10	US-09-731-816-15	Sequence 15, Appl
c 31	11	55.0	86	10	US-09-768-826-3	Sequence 3, Appli
c 32	11	55.0	86	10	US-09-726-348-6	Sequence 6, Appli
c 33	11	55.0	86	10	US-09-244-694-11	Sequence 11, Appl
c 34	11	55.0	86	10	US-09-904-615-3	Sequence 3, Appli
c 35	11	55.0	86	10	US-09-726-643-3	Sequence 3, Appli
c 36	11	55.0	86	10	US-09-764-887-3	Sequence 3, Appli
c 37	11	55.0	86	10	US-09-764-870-3	Sequence 3, Appli
c 38	11	55.0	86	10	US-09-925-302-889	Sequence 889, App
c 39	11	55.0	86	10	US-09-908-711-3	Sequence 3, Appli
c 40	11	55.0	86	10	US-09-779-879A-12	Sequence 12, Appl
c 41	11	55.0	86	10	US-09-925-301-1687	Sequence 1687, Ap
c 42	11	55.0	86	10	US-09-908-599-11	Sequence 11, Appl
c 43	11	55.0	86	10	US-09-925-299-1549	Sequence 1549, Ap
c 44	11	55.0	86	10	US-09-814-122-3	Sequence 3, Appli
c 45	11	55.0	86	10	US-09-764-869-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-09-921-942B-2
; Sequence 2, Application US/09921942B
; Patent No. US20020127724A1
; GENERAL INFORMATION:
; APPLICANT: Battelle Memorial Institute
; TITLE OF INVENTION: Cloning of a Yeast Alpha-Amylase Promoter and Its Regulated He
; TITLE OF INVENTION: Expression
; FILE REFERENCE: E-1824 CIP
; CURRENT APPLICATION NUMBER: US/09/921,942B
; CURRENT FILING DATE: 2002-02-13
; PRIOR APPLICATION NUMBER: 09/632,313
; PRIOR FILING DATE: 2000-08-04
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 33
; TYPE: DNA
; ORGANISM: artificial sequence
; FEATURE:
; OTHER INFORMATION: oligonucleotide primer
US-09-921-942B-2

Query Match 66.0%; Score 13.2; DB 10; Length 33;
Best Local Similarity 83.3%; Pred. No. 2.7e+02;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy       3 GGACAAATCGAGTTATCA 20
           | || |||||| ||||||
Db       8 GCACCAATCGATTTATCA 25
```

RESULT 2
US-09-100-812-4
; Sequence 4, Application US/09100812
; Patent No. US20020168707A1
; GENERAL INFORMATION:
; APPLICANT: Graham, Michael W.
; TITLE OF INVENTION: SYNTHETIC GENES AND GENETIC CONSTRUCTS COMPRISING SAME I
; FILE REFERENCE: 11535
; CURRENT APPLICATION NUMBER: US/09/100,812
; CURRENT FILING DATE: 2002-07-02

```
;  NUMBER OF SEQ ID NOS: 10
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
;   LENGTH: 28
;   TYPE: DNA
;   ORGANISM: Artificial sequence
;   FEATURE:
;   OTHER INFORMATION: Primer Bev-4
US-09-100-812-4

  Query Match            65.0%;  Score 13;  DB 9;  Length 28;
  Best Local Similarity  100.0%;  Pred. No. 3.4e+02;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3 GGACAAATCGAGT 15
           |||||||||||||
Db      11 GGACAAATCGAGT 23


RESULT 3
US-09-100-812-1
; Sequence 1, Application US/09100812
; Patent No. US20020168707A1
; GENERAL INFORMATION:
;  APPLICANT: Graham, Michael W.
;  TITLE OF INVENTION: SYNTHETIC GENES AND GENETIC CONSTRUCTS COMPRISING SAME I
;  FILE REFERENCE: 11535
;  CURRENT APPLICATION NUMBER: US/09/100,812
;  CURRENT FILING DATE:  2002-07-02
;  NUMBER OF SEQ ID NOS: 10
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
;   LENGTH: 38
;   TYPE: DNA
;   ORGANISM: Artificial sequence
;   FEATURE:
;   OTHER INFORMATION: Primer Bev-1
US-09-100-812-1

  Query Match            65.0%;  Score 13;  DB 9;  Length 38;
  Best Local Similarity  100.0%;  Pred. No. 3.5e+02;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3 GGACAAATCGAGT 15
           |||||||||||||
Db      21 GGACAAATCGAGT 33


RESULT 4
US-09-923-876-827
; Sequence 827, Application US/09923876
; Patent No. US20020013958A1
; GENERAL INFORMATION:
;  APPLICANT: Lalgudi, Raghunath V.
;  APPLICANT:  Kamigaki, Laura Y. (Ito)
;  APPLICANT:  Sherman, Bradley K.
;  TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN SEEDLING
;  FILE REFERENCE: PL-0012-1 CON
;  CURRENT APPLICATION NUMBER: US/09/923,876
;  CURRENT FILING DATE:  2001-08-06
;  PRIOR APPLICATION NUMBER: 09/298,329
;  PRIOR FILING DATE: 1999-04-21
;  PRIOR APPLICATION NUMBER: 60/085,331
;  PRIOR FILING DATE: 1998-05-05
;  NUMBER OF SEQ ID NOS: 6332
;  SOFTWARE: PERL Program
; SEQ ID NO 827
;   LENGTH: 96
;   TYPE: DNA
;   ORGANISM: Zea mays
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. US20020013958A1 700157647H1
;   NAME/KEY: unsure
;   LOCATION: 25, 79, 81, 84-87, 91-92
;   OTHER INFORMATION: a, t, c, g, or other
US-09-923-876-827

  Query Match            61.0%;  Score 12.2;  DB 10;  Length 96;
  Best Local Similarity  82.4%;  Pred. No. 1e+03;
  Matches   14;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 CGGACAAATCGAGTTAT 18
           || ||||||||||  ||
Db      43 CGCACAAATCGAGAAAT 59


RESULT 5
US-09-801-274-833
; Sequence 833, Application US/09801274
; Patent No. US20020032319A1
; GENERAL INFORMATION:
;  APPLICANT: Cargill, Michele
;  APPLICANT:  Ireland, James S.
;  APPLICANT:  Lander, Eric S.
;  TITLE OF INVENTION: HUMAN SINGLE NUCLEOTIDE POLYMORPHISMS
;  FILE REFERENCE: 2825.2009-001
;  CURRENT APPLICATION NUMBER: US/09/801,274
;  CURRENT FILING DATE:  2001-03-07
;  PRIOR APPLICATION NUMBER: US 60/187,510
;  PRIOR FILING DATE: 2000-03-07
;  PRIOR APPLICATION NUMBER: US 60/206,129
;  PRIOR FILING DATE: 2000-05-22
;  NUMBER OF SEQ ID NOS: 1802
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 833
;   LENGTH: 31
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-801-274-833

  Query Match            58.0%;  Score 11.6;  DB 10;  Length 31;
  Best Local Similarity  70.0%;  Pred. No. 2e+03;
  Matches   14;  Conservative    1;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1 TCGGACAAATCGAGTTATCA 20
           |||||| | || ||: | ||
Db       2 TCGGACTACTCCAGKGAGCA 21


RESULT 6
US-10-115-701A-15/c
; Sequence 15, Application US/10115701A
; Patent No. US20020155996A1
; GENERAL INFORMATION:
;  APPLICANT: Murgita, Robert A.
;  TITLE OF INVENTION: Recombinant Alpha-Fetoprotein for
;  TITLE OF INVENTION:  Treating and Diagnosing Cancers
;  FILE REFERENCE: 06727/004003
;  CURRENT APPLICATION NUMBER: US/10/115,701A
;  CURRENT FILING DATE:  2002-04-04
;  PRIOR APPLICATION NUMBER: 08/758,757
;  PRIOR FILING DATE: 2000-11-01
;  PRIOR APPLICATION NUMBER: 08/758,757
;  PRIOR FILING DATE: 1998-04-09
;  PRIOR APPLICATION NUMBER: 08/758,757
;  PRIOR FILING DATE: 1996-12-03
;  PRIOR APPLICATION NUMBER: 08/377,311
;  PRIOR FILING DATE: 1995-01-24
;  NUMBER OF SEQ ID NOS: 16
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 15
;   LENGTH: 33
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-115-701A-15
```

```
  Query Match             58.0%;  Score 11.6;  DB 9;  Length 33;
  Best Local Similarity   77.8%;  Pred. No. 2e+03;
  Matches   14;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 TCGGACAAATCGAGTTAT 18
          | ||||| ||||| || |
Db     20 TAGGACATATCGATTTTT 3


RESULT 7
US-09-294-093B-1011
; Sequence 1011, Application US/09294093B
; Patent No. US20010051335A1
; GENERAL INFORMATION:
;  APPLICANT:      Lalgudi, Raghunath, V.
;  APPLICANT:  Ito, Laura, Y.
;  APPLICANT:  Sherman, Bradley, K.
;  TITLE OF INVENTION:  POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN TASSEL
;  FILE REFERENCE:  PL-0009 US
;  CURRENT APPLICATION NUMBER: US/09/294,093B
;  CURRENT FILING DATE:  1999-04-16
;  PRIOR APPLICATION NUMBER:  60/082,567
;  PRIOR FILING DATE:  April 21, 1998
;  NUMBER OF SEQ ID NOS:      6207
;  SOFTWARE:      PERL Program
; SEQ ID NO 1011
;   LENGTH: 91
;   TYPE: DNA
;   ORGANISM: Zea mays
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. US20010051335A1 700343643H1
;   NAME/KEY: unsure
;   LOCATION: 43, 67
;   OTHER INFORMATION: a, t, c, g, or other
US-09-294-093B-1011

  Query Match             58.0%;  Score 11.6;  DB 10;  Length 91;
  Best Local Similarity   73.7%;  Pred. No. 2.2e+03;
  Matches   14;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      1 TCGGACAAATCGAGTTATC 19
          ||| ||| |   |||||||
Db     58 TCGCACACANAAAGTTATC 76


RESULT 8
US-09-466-035-22
; Sequence 22, Application US/09466035
; Patent No. US20020165172A1
;   GENERAL INFORMATION:
;        APPLICANT: SALLBERG, MATTI
;                   MILICH, DAVID R.
;                   LEE, WILLIAM T.L.
;        TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
;                            INTRACELLULAR DISEASES
;        NUMBER OF SEQUENCES: 86
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Robins & Pasternak LLP
;             STREET: 545 Middlefield Road, Suite 180
;             CITY: Menlo Park
;             STATE: California
;             COUNTRY: U.S.
;             ZIP: 94025
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.25
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/466,035
;             FILING DATE: 17-Dec-1999
;             CLASSIFICATION: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Pasternak, Dahna S.
;             REGISTRATION NUMBER: 41,411
;             REFERENCE/DOCKET NUMBER: 2300-1231.01
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 650-325-7812
;             TELEFAX: 650-325-7823
;             TELEX: <Unknown>
;   INFORMATION FOR SEQ ID NO: 22:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 35 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        SEQUENCE DESCRIPTION: SEQ ID NO: 22:
US-09-466-035-22

  Query Match             57.0%;  Score 11.4;  DB 9;  Length 35;
  Best Local Similarity   92.3%;  Pred. No. 2.5e+03;
  Matches   12;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      8 AATCGAGTTATCA 20
          | ||||||||||||
Db      2 ACTCGAGTTATCA 14


RESULT 9
US-09-912-679-22
; Sequence 22, Application US/09912679
; Patent No. US20020141974A1
;   GENERAL INFORMATION:
;        APPLICANT: Jolly, Douglas J.
;                   Chang, Stephen M.W.
;                   Lee, William T.L.
;                   Townsend, Kay
;                   O'Dea, Joanne
;        TITLE OF INVENTION: HEPATITIS THERAPEUTICS
;        NUMBER OF SEQUENCES: 84
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Seed and Berry
;             STREET: 6300 Columbia Center, 701 Fifth Avenue
;             CITY: Seattle
;             STATE: Washington
;             COUNTRY: U.S.
;             ZIP: 98104
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.25
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/912,679
;             FILING DATE: 07-Jun-1995
;             CLASSIFICATION: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: McMasters, David D.
;             REGISTRATION NUMBER: 33,963
;             REFERENCE/DOCKET NUMBER: 930049.407C5
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 206-622-4900
;             TELEFAX: 206-682-6031
;             TELEX: 3723836
;   INFORMATION FOR SEQ ID NO: 22:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 35 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        SEQUENCE DESCRIPTION: SEQ ID NO: 22:
US-09-912-679-22

  Query Match             57.0%;  Score 11.4;  DB 10;  Length 35;
```

```
  Best Local Similarity   92.3%;  Pred. No. 2.5e+03;
  Matches   12;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       8 AATCGAGTTATCA 20
           | ||||||||||||
Db       2 ACTCGAGTTATCA 14


RESULT 10
US-09-978-295A-181
; Sequence 181, Application US/09978295A
; Patent No. US20020156006A1
; GENERAL INFORMATION:
;  APPLICANT: Ashkenazi, Avi
;  APPLICANT:  Baker Kevin P.
;  APPLICANT:  Botstein, David
;  APPLICANT:  Desnoyers, Luc
;  APPLICANT:  Eaton, Dan
;  APPLICANT:  Ferrara, Napoleon
;  APPLICANT:  Filvaroff, Ellen
;  APPLICANT:  Fong, Sherman
;  APPLICANT:  Gao, Wei-Qiang
;  APPLICANT:  Gerber, Hanspeter
;  APPLICANT:  Gerritsen, Mary E.
;  APPLICANT:  Goddard, Audrey
;  APPLICANT:  Godowski, Paul J.
;  APPLICANT:  Grimaldi, J. Christopher
;  APPLICANT:  Gurney, Austin L.
;  APPLICANT:  Hillan, Kenneth J
;  APPLICANT:  Kljavin, Ivar J.
;  APPLICANT:  Kuo, Sophia S.
;  APPLICANT:  Napier, Mary A.
;  APPLICANT:  Pan, James;
;  APPLICANT:  Paoni, Nicholas F.
;  APPLICANT:  Roy, Margaret Ann
;  APPLICANT:  Shelton, David L.
;  APPLICANT:  Stewart, Timothy A.
;  APPLICANT:  Tumas, Daniel
;  APPLICANT:  Williams, P. Mickey
;  APPLICANT:  Wood, William I.
;  TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
;  TITLE OF INVENTION:  Acids Encoding the Same
;  FILE REFERENCE: P2630P1C11
;  CURRENT APPLICATION NUMBER: US/09/978,295A
;  CURRENT FILING DATE:  2001-10-15
;  PRIOR APPLICATION NUMBER: 09/918585
;  PRIOR FILING DATE: 2001-07-30
;  PRIOR APPLICATION NUMBER: 60/062250
;  PRIOR FILING DATE: 1997-10-17
;  PRIOR APPLICATION NUMBER: 60/064249
;  PRIOR FILING DATE: 1997-11-03
;  PRIOR APPLICATION NUMBER: 60/065311
;  PRIOR FILING DATE: 1997-11-13
;  PRIOR APPLICATION NUMBER: 60/066364
;  PRIOR FILING DATE: 1997-11-21
;  PRIOR APPLICATION NUMBER: 60/077450
;  PRIOR FILING DATE: 1998-03-10
;  PRIOR APPLICATION NUMBER: 60/077632
;  PRIOR FILING DATE: 1998-03-11
;  PRIOR APPLICATION NUMBER: 60/077641
;  PRIOR FILING DATE: 1998-03-11
;  PRIOR APPLICATION NUMBER: 60/077649
;  PRIOR FILING DATE: 1998-03-11
;  PRIOR APPLICATION NUMBER: 60/077791
;  PRIOR FILING DATE: 1998-03-12
;  PRIOR APPLICATION NUMBER: 60/078004
;  PRIOR FILING DATE: 1998-03-13
;  PRIOR APPLICATION NUMBER: 60/078886
;  PRIOR FILING DATE: 1998-03-20
;  PRIOR APPLICATION NUMBER: 60/078936
;  PRIOR FILING DATE: 1998-03-20
;  PRIOR APPLICATION NUMBER: 60/078910
;  PRIOR FILING DATE: 1998-03-20
;  PRIOR APPLICATION NUMBER: 60/078939
;  PRIOR FILING DATE: 1998-03-20
;  PRIOR APPLICATION NUMBER: 60/079294
;  PRIOR FILING DATE: 1998-03-25
;  PRIOR APPLICATION NUMBER: 60/079656
;  PRIOR FILING DATE: 1998-03-26
;  PRIOR APPLICATION NUMBER: 60/079664
;  PRIOR FILING DATE: 1998-03-27
;  PRIOR APPLICATION NUMBER: 60/079689
;  PRIOR FILING DATE: 1998-03-27
;  PRIOR APPLICATION NUMBER: 60/079663
;  PRIOR FILING DATE: 1998-03-27
;  PRIOR APPLICATION NUMBER: 60/079728
;  PRIOR FILING DATE: 1998-03-27
;  PRIOR APPLICATION NUMBER: 60/079786
;  PRIOR FILING DATE: 1998-03-27
;  PRIOR APPLICATION NUMBER: 60/079920
;  PRIOR FILING DATE: 1998-03-30
;  PRIOR APPLICATION NUMBER: 60/079923
;  PRIOR FILING DATE: 1998-03-30
;  PRIOR APPLICATION NUMBER: 60/080105
;  PRIOR FILING DATE: 1998-03-31
;  PRIOR APPLICATION NUMBER: 60/080107
;  PRIOR FILING DATE: 1998-03-31
;  PRIOR APPLICATION NUMBER: 60/080165
;  PRIOR FILING DATE: 1998-03-31
;  PRIOR APPLICATION NUMBER: 60/080194
;  PRIOR FILING DATE: 1998-03-31
;  PRIOR APPLICATION NUMBER: 60/080327
;  PRIOR FILING DATE: 1998-04-01
;  PRIOR APPLICATION NUMBER: 60/080328
;  PRIOR FILING DATE: 1998-04-01
;  PRIOR APPLICATION NUMBER: 60/080333
;  PRIOR FILING DATE: 1998-04-01
;  PRIOR APPLICATION NUMBER: 60/080334
;  PRIOR FILING DATE: 1998-04-01
;  PRIOR APPLICATION NUMBER: 60/081070
;  PRIOR FILING DATE: 1998-04-08
;  PRIOR APPLICATION NUMBER: 60/081049
;  PRIOR FILING DATE: 1998-04-08
;  PRIOR APPLICATION NUMBER: 60/081071
;  PRIOR FILING DATE: 1998-04-08
;  PRIOR APPLICATION NUMBER: 60/081195
;  PRIOR FILING DATE: 1998-04-08
;  PRIOR APPLICATION NUMBER: 60/081203
;  PRIOR FILING DATE: 1998-04-09
;  PRIOR APPLICATION NUMBER: 60/081229
;  PRIOR FILING DATE: 1998-04-09
;  PRIOR APPLICATION NUMBER: 60/081955
;  PRIOR FILING DATE: 1998-04-15
;  PRIOR APPLICATION NUMBER: 60/081817
;  PRIOR FILING DATE: 1998-04-15
;  PRIOR APPLICATION NUMBER: 60/081819
;  PRIOR FILING DATE: 1998-04-15
;  PRIOR APPLICATION NUMBER: 60/081952
;  PRIOR FILING DATE: 1998-04-15
;  PRIOR APPLICATION NUMBER: 60/081838
;  PRIOR FILING DATE: 1998-04-15
;  PRIOR APPLICATION NUMBER: 60/082568
;  PRIOR FILING DATE: 1998-04-21
;  PRIOR APPLICATION NUMBER: 60/082569
;  PRIOR FILING DATE: 1998-04-21
;  PRIOR APPLICATION NUMBER: 60/082704
;  PRIOR FILING DATE: 1998-04-22
;  PRIOR APPLICATION NUMBER: 60/082804
;  PRIOR FILING DATE: 1998-04-22
;  PRIOR APPLICATION NUMBER: 60/082700
;  PRIOR FILING DATE: 1998-04-22
;  PRIOR APPLICATION NUMBER: 60/082797
;  PRIOR FILING DATE: 1998-04-22
;  PRIOR APPLICATION NUMBER: 60/082796
;  PRIOR FILING DATE: 1998-04-23
;  PRIOR APPLICATION NUMBER: 60/083336
```

; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/083392
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083495
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083496
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083499
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083545
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083554
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083558
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083559
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083500
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083742
; PRIOR FILING DATE: 1998-04-30
; PRIOR APPLICATION NUMBER: 60/084366
; PRIOR FILING DATE: 1998-05-05
; PRIOR APPLICATION NUMBER: 60/084414
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084441
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084637
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084639
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084640
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084598
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084627
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084643
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/085339
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085338
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085323
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085582
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085700
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085689
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085580
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085573
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085704
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085697

Query Match 56.0%; Score 11.2; DB 9; Length 44;
Best Local Similarity 81.2%; Pred. No. 3.3e+03;
Matches 13; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy       3 GGACAAATCGAGTTAT 18
           ||||||  |||||| |
Db      20 GGACAAGCCGAGTTTT 35
```

RESULT 11
US-09-978-697-181
; Sequence 181, Application US/09978697
; Patent No. US20020169284A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleon
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Kuo, Sophia S.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James;
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Shelton, David L.
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2630P1C27
; CURRENT APPLICATION NUMBER: US/09/978,697
; CURRENT FILING DATE: 2001-10-16
; PRIOR APPLICATION NUMBER: 09/918585
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/064249
; PRIOR FILING DATE: 1997-11-03
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066364
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: 60/077450
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: 60/077632
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077641
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077649
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077791
; PRIOR FILING DATE: 1998-03-12
; PRIOR APPLICATION NUMBER: 60/078004
; PRIOR FILING DATE: 1998-03-13
; PRIOR APPLICATION NUMBER: 60/078886
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078936
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078939
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079664

; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079663
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079786
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079920
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 60/079923
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 60/080105
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080107
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080165
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080194
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080327
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080328
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080333
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080334
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/081070
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081049
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081071
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081195
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081203
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081229
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081955
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081817
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081819
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081952
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081838
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/082568
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082569
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082704
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082804
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082700
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082797
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082796
; PRIOR FILING DATE: 1998-04-23
; PRIOR APPLICATION NUMBER: 60/083336
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/083392
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083495
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083496
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083499
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083545
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083554
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083558
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083559
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083500
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083742
; PRIOR FILING DATE: 1998-04-30
; PRIOR APPLICATION NUMBER: 60/084366
; PRIOR FILING DATE: 1998-05-05
; PRIOR APPLICATION NUMBER: 60/084414
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084441
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084637
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084639
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084640
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084598
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-5-07
; PRIOR APPLICATION NUMBER: 60/084627
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084643
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/085339
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085338
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085323
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085582
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085700
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085689
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085580
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085573
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085704
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085697

Query Match 56.0%; Score 11.2; DB 9; Length 44;
Best Local Similarity 81.2%; Pred. No. 3.3e+03;
Matches 13; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy       3 GGACAAATCGAGTTAT 18
           ||||||  |||||| |
Db      20 GGACAAGCCGAGTTTT 35
```

RESULT 12
US-09-923-876-3345
; Sequence 3345, Application US/09923876
; Patent No. US20020013958A1
; GENERAL INFORMATION:
; APPLICANT: Lalgudi, Raghunath V.

```
;  APPLICANT:  Kamigaki, Laura Y. (Ito)
;  APPLICANT:  Sherman, Bradley K.
;  TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN SEEDLING
;  FILE REFERENCE: PL-0012-1 CON
;  CURRENT APPLICATION NUMBER: US/09/923,876
;  CURRENT FILING DATE:  2001-08-06
;  PRIOR APPLICATION NUMBER: 09/298,329
;  PRIOR FILING DATE: 1999-04-21
;  PRIOR APPLICATION NUMBER: 60/085,331
;  PRIOR FILING DATE: 1998-05-05
;  NUMBER OF SEQ ID NOS: 6332
;  SOFTWARE: PERL Program
; SEQ ID NO 3345
;   LENGTH: 77
;   TYPE: DNA
;   ORGANISM: Zea mays
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No. US20020013958A1 700162067H1
;   NAME/KEY: unsure
;   LOCATION: 6, 61, 74, 76
;   OTHER INFORMATION: a, t, c, g, or other
US-09-923-876-3345

  Query Match            56.0%;  Score 11.2;  DB 10;  Length 77;
  Best Local Similarity  81.2%;  Pred. No. 3.6e+03;
  Matches   13;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       5 ACAAATCGAGTTATCA 20
           ||||||| | |||| |
Db      27 ACAAATCCAATTATGA 42


RESULT 13
US-09-878-574-11180
; Sequence 11180, Application US/09878574
; Patent No. US20020110548A1
; GENERAL INFORMATION:
;  APPLICANT: Byrum, Joseph R.
;  APPLICANT:  La Rosa, Thomas J.
;  APPLICANT:  Thompson, Michael D.
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
;  TITLE OF INVENTION: Plants
;  FILE REFERENCE: 38-21(15401)B
;  CURRENT APPLICATION NUMBER: US/09/878,574
;  CURRENT FILING DATE:  2001-12-21
;  PRIOR APPLICATION NUMBER: 09/333,535
;  PRIOR FILING DATE: 1999-06-14
;  NUMBER OF SEQ ID NOS: 15775
; SEQ ID NO 11180
;   LENGTH: 97
;   TYPE: DNA
;   ORGANISM: Glycine max
;   OTHER INFORMATION: Clone ID: 701064141H1
US-09-878-574-11180

  Query Match            56.0%;  Score 11.2;  DB 10;  Length 97;
  Best Local Similarity  81.2%;  Pred. No. 3.6e+03;
  Matches   13;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 CGGACAAATCGAGTTA 17
           || || | ||||||||
Db      34 CGAACGAGTCGAGTTA 49


RESULT 14
US-09-864-761-32354
; Sequence 32354, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 32354
;   LENGTH: 99
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AL159165.1
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 9.7
;   OTHER INFORMATION: NT HIT: Z72955.1, EVALUE 5.30e-02
US-09-864-761-32354

  Query Match            56.0%;  Score 11.2;  DB 10;  Length 99;
  Best Local Similarity  81.2%;  Pred. No. 3.6e+03;
  Matches   13;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       4 GACAAATCGAGTTATC 19
           ||||||| ||| | ||
Db      31 GACAAATAGAGGTTTC 46


RESULT 15
US-09-971-309-85
; Sequence 85, Application US/09971309
; Patent No. US20020106675A1
; GENERAL INFORMATION:
;  APPLICANT: UEMORI, Takashi
;  APPLICANT:  SATO, Yoshimi
;  APPLICANT:  FUJITA, Tomoko
;  APPLICANT:  MIYAKE, Kazue
;  APPLICANT:  MUKAI, Hiroyuki
```

```
;  APPLICANT:  ASADA, Kiyozo
;  APPLICANT:  KATO, Ikunoshin
;  TITLE OF INVENTION: DNA POLYMERASE-RELATED FACTORS
;  FILE REFERENCE: 1422-0494P
;  CURRENT APPLICATION NUMBER: US/09/971,309
;  CURRENT FILING DATE:  2001-10-05
;  PRIOR APPLICATION NUMBER: US 09/446,504
;  PRIOR FILING DATE: 1999-12-23
;  PRIOR APPLICATION NUMBER: PCT/JP98/02845
;  PRIOR FILING DATE: 1998-06-24
;  PRIOR APPLICATION NUMBER: JP 9-187496
;  PRIOR FILING DATE: 1997-06-26
;  PRIOR APPLICATION NUMBER: JP 9-320692
;  PRIOR FILING DATE: 1997-11-27
;  NUMBER OF SEQ ID NOS: 92
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 85
;   LENGTH: 35
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Synthetic DNA
US-09-971-309-85

  Query Match             55.0%;  Score 11;  DB 10;  Length 35;
  Best Local Similarity   73.7%;  Pred. No. 4.2e+03;
  Matches   14;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       2 CGGACAAATCGAGTTATCA 20
           || |||| | | || ||||
Db      15 CGTACAACTGGCGTAATCA 33


Search completed: November 23, 2002, 07:10:40
Job time : 17.8 secs
```

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 26, 2002, 04:16:10 ; Search time 798.5 Seconds
(without alignments)
405.648 Million cell updates/sec

Title: US-09-296-264-23
Perfect score: 20
Sequence: 1 tcggacaaatcgagttatca 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 357874

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	15.2	76.0	62	14	C57360	C57360 C57360 Yuji
c	2	13.8	69.0	64	10	BE239292	BE239292 SWOvL2CAS
	3	13.8	69.0	98	17	AZ308835	AZ308835 1M0012E17
	4	13.6	68.0	85	12	BG236544	BG236544 nai45a06.
	5	13.2	66.0	70	17	AZ877648	AZ877648 BG00642-3
	6	13.2	66.0	76	9	AI222966	AI222966 qh23e04.x
c	7	13.2	66.0	100	9	AA509111	AA509111 MBAFCX7D0
	8	12.8	64.0	79	17	BH812020	BH812020 SALK_0610
	9	12.6	63.0	62	17	AZ661477	AZ661477 1M0540A14
c	10	12.6	63.0	75	9	AU031313	AU031313 AU031313
c	11	12.6	63.0	97	17	AL757650	AL757650 Arabidops
c	12	12.6	63.0	100	9	AA671696	AA671696 vl09b07.r
	13	12.4	62.0	55	17	TA83D07Q	AL462388 T. brucei
c	14	12.4	62.0	57	10	AV832570	AV832570 AV832570
	15	12.4	62.0	84	17	AQ073044	AQ073044 EP(3)3592
c	16	12.4	62.0	99	13	BI699914	BI699914 sag49e12.
	17	12.4	62.0	100	9	AL797747	AL797747 AL797747
	18	12.2	61.0	32	9	AU013431	AU013431 AU013431
	19	12.2	61.0	58	9	AU008450	AU008450 AU008450
c	20	12.2	61.0	67	17	TA282H03Q	AL485604 T. brucei
c	21	12.2	61.0	72	17	AZ765633	AZ765633 1M0562M11
c	22	12.2	61.0	80	17	AQ026360	AQ026360 l(3)rK315
	23	12.2	61.0	85	14	W39283	W39283 zc76f01.r1
c	24	12.2	61.0	87	12	BF507305	BF507305 8949P-10
c	25	12.2	61.0	92	9	AU007487	AU007487 AU007487
c	26	12.2	61.0	100	14	BM844245	BM844245 K-EST0122
	27	12.2	61.0	100	14	BQ972786	BQ972786 QHI10I05.
c	28	12	60.0	54	17	BH847390	BH847390 SALK_0531
c	29	12	60.0	64	10	AW713312	AW713312 g7b05ne.f
c	30	12	60.0	64	17	BH846663	BH846663 SALK_0096
c	31	12	60.0	68	17	HSMC23D06	X88266 H.sapiens D
c	32	12	60.0	71	17	BH890326	BH890326 3526_1_13
	33	12	60.0	74	17	AZ607963	AZ607963 1M0430M12
	34	12	60.0	81	10	AV958150	AV958150 AV958150
c	35	12	60.0	82	12	BG315433	BG315433 PO2.0.325
c	36	12	60.0	82	17	BH796052	BH796052 1008092E1
	37	12	60.0	83	14	BQ482609	BQ482609 ke51g10.y
	38	12	60.0	93	17	AZ657428	AZ657428 1M0533B09
c	39	12	60.0	95	17	BH636055	BH636055 1008008G0
c	40	12	60.0	97	9	AA630415	AA630415 ac09c07.s
c	41	12	60.0	98	10	AV954357	AV954357 AV954357
	42	12	60.0	100	9	AI202525	AI202525 qs71g08.x
	43	12	60.0	100	13	BG996454	BG996454 CM0-HT129
	44	11.8	59.0	64	10	AW600227	AW600227 SWOvL3CAN
	45	11.8	59.0	64	12	BF199405	BF199405 SWOv3MCAM

ALIGNMENTS

RESULT 1
C57360
LOCUS C57360 62 bp mRNA linear EST 22-SEP-1997
DEFINITION C57360 Yuji Kohara unpublished cDNA Caenorhabditis elegans cDNA clone yk302f12 3', mRNA sequence.
ACCESSION C57360
VERSION C57360.1 GI:2416065
KEYWORDS EST.
SOURCE Caenorhabditis elegans.
ORGANISM Caenorhabditis elegans
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea ; Rhabditidae; Peloderinae; Caenorhabditis.
REFERENCE 1 (bases 1 to 62)
AUTHORS Kohara,Y., Motohashi,T., Tabara,H., Watanabe,H., Sugimoto,A., Sano ,M., Miyata,A. and Nishigaki,A.
TITLE Expression map of the C.elegans genome
JOURNAL Unpublished (1996)
COMMENT Contact: Yuji Kohara
Genome Biology Lab.
National Institute of Genetics
Yata 1111, Mishima, Shizuoka 411, Japan
Tel: 81-559-81-6854
Fax: 81-559-81-6855
Email: ykohara@lab.nig.ac.jp.
FEATURES Location/Qualifiers
source 1. .62
/organism="Caenorhabditis elegans"
/strain="CB1489 him-8(e1489)"
/db_xref="taxon:6239"

```
                     /clone="yk302f12"
                     /clone_lib="Yuji Kohara unpublished cDNA"
                     /sex="hermaphrodite, male"
                     /tissue_type="whole animal"
                     /dev_stage="varied"
BASE COUNT        9 a     20 c     14 g     19 t
ORIGIN

  Query Match           76.0%;  Score 15.2;  DB 14;  Length 62;
  Best Local Similarity   85.0%;  Pred. No. 5e+02;
  Matches   17;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 TCGGACAAATCGAGTTATCA 20
          ||||||||||||  || |||
Db      2 TCGGACAAATCGTCTTGTCA 21


RESULT 2
BE239292/c
LOCUS       BE239292                64 bp    mRNA    linear   EST 11-JUL-2000
DEFINITION  SWOvL2CAS08H02SK Onchocerca volvulus L2 larvae cDNA (SAW98MLW-OvL2)
            Onchocerca volvulus cDNA clone SWOvL2CAS08H02 5', mRNA sequence.
ACCESSION   BE239292
VERSION     BE239292.1  GI:9034256
KEYWORDS    EST.
SOURCE      Onchocerca volvulus.
  ORGANISM  Onchocerca volvulus
            Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
            Onchocercidae; Onchocerca.
REFERENCE   1  (bases 1 to 64)
  AUTHORS   Williams,S.A.
  TITLE     Genes expressed in L2 larvae of Onchocerca volvulus
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Steven A. Williams
            Molecular Parasitology
            Smith College Department of Biological Sciences
            Department of Biological Sciences, Clark Science Center, Smith
            College, Northampton, MA, 01063, USA
            Tel: 4135853826
            Fax: 4135853786
            Email: genome@smith.edu
            Seq primer: pBluescript SK.
FEATURES             Location/Qualifiers
     source          1. .64
                     /organism="Onchocerca volvulus"
                     /db_xref="taxon:6282"
                     /clone="SWOvL2CAS08H02"
                     /clone_lib="Onchocerca volvulus L2 larvae cDNA
                     (SAW98MLW-OvL2)"
                     /dev_stage="L2"
                     /lab_host="XL1-Blue MRF'"
                     /note="Vector: Lambda Uni-ZAP XR; Site_1: Eco RI; Site_2:
                     Xho I; Filarial nematode parasite of humans. mRNA was
                     prepared from approximately 9,000 L2s isolated from
                     infected  mosquitoes from Kumba, Cameroon and converted to
                     double-stranded cDNA using  reverse transcriptase and
                     oligo(dT) followed by RNase H and DNA pol I. The  library
                     has 7.3 x 10E4 independent recombinants and the average
                     insert size is  approximately 1kb.The library was
                     constructed by Michelle Lizotte-Waniewski.  The library is
                     available from Dr.S.A.Williams, email: genome@smith.edu."
BASE COUNT       21 a     13 c     12 g     18 t
ORIGIN

  Query Match           69.0%;  Score 13.8;  DB 10;  Length 64;
  Best Local Similarity   88.2%;  Pred. No. 2.9e+03;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 TCGGACAAATCGAGTTA 17
          ||| ||||||||||| |
Db     57 TCGTACAAATCGAGTCA 41
```

```
RESULT 3
AZ308835
LOCUS       AZ308835                98 bp    DNA     linear   GSS 29-SEP-2000
DEFINITION  1M0012E17F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0012E17 F, DNA sequence.
ACCESSION   AZ308835
VERSION     AZ308835.1  GI:10349224
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 98)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0012  row: E  column: 17
            Seq primer: CGTTGTAAAACGACGGCCAGT
            Class: plasmid ends
            High quality sequence stop: 98.
FEATURES             Location/Qualifiers
     source          1. .98
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC1M0012E17"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
BASE COUNT       29 a     13 c     16 g     40 t
ORIGIN

  Query Match           69.0%;  Score 13.8;  DB 17;  Length 98;
  Best Local Similarity   88.2%;  Pred. No. 3.6e+03;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      4 GACAAATCGAGTTATCA 20
          |||| |||||||| ||
Db     56 GACACATCGAGTTAGCA 72
```

```
RESULT 4
BG236544
LOCUS       BG236544                 85 bp    mRNA    linear   EST 12-FEB-2001
DEFINITION  na145a06.x1 NCI_CGAP_HN20 Homo sapiens cDNA clone IMAGE:4263058 3',
            mRNA sequence.
ACCESSION   BG236544
VERSION     BG236544.1  GI:12750391
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 85)
  AUTHORS   NCI/NIDR-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute / National Institute of Dental Research,
            Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            cDNA Library Preparation: David B. Krizman, Ph.D.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL, send email to:
            info@image.llnl.gov
            Seq primer: -40UP from Gibco.
FEATURES             Location/Qualifiers
     source          1. .85
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4263058"
                     /clone_lib="NCI_CGAP_HN20"
                     /lab_host="DH10B"
                     /note="Organ: normal head/neck tissue; Vector: pAMP1; mRNA
                     made from head/neck tissue, cDNA made by oligo-dT
                     priming. Directionally cloned into UDG sites.
                     Size-selected on  agarose gel, average insert size 300 bp.
                     Primary library.  cDNA Library Preparation: David B.
                     Krizman, Ph.D."
BASE COUNT       22 a     15 c     13 g     33 t      2 others
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 12;  Length 85;
  Best Local Similarity   80.0%;  Pred. No. 4.3e+03;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 TCGGACAAATCGAGTTATCA 20
           | | |||||| || ||||||
Db      25 TTGTACAAATTGAATTATCA 44


RESULT 5
AZ877648
LOCUS       AZ877648                 70 bp    DNA     linear   GSS 26-FEB-2001
DEFINITION  BG00642-3prime Drosophila melanogaster P{GT1} P element insertion
            lines Drosophila melanogaster genomic Sequence recovered from 3'
            end of P element, DNA sequence.
ACCESSION   AZ877648
VERSION     AZ877648.1  GI:13124934
KEYWORDS    GSS.
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 70)
  AUTHORS   Levis,R., Hoskins,R., Liao,G., Mozden,N., Tsang,G., He,Y., Karpen
            ,G., Bellen,H., Rubin,G. and Spradling,A.
  TITLE     The Berkeley Drosophila Genome Project Gene Disruption Project
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Gerald Rubin
            Berkeley Drosophila Genome Project
            University of California, Berkeley
            LSA Building, Berkeley, CA 94720-3200, USA
            Fax: 5106439947
            Email: gerry@fruitfly.berkeley.edu
            Sequence recovery method was inverse PCR.
            Sequence orientation is forward strand relative to 5' end of P
            element
             The P element insertion position is base 1 in the 70 bases. This
            insertion position refers to the first base of the 8 base target
            recognition sequence.
            Class: transposon-tagged.
FEATURES             Location/Qualifiers
     source          1. .70
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone_lib="Drosophila melanogaster P{GT1} P element
                     insertion lines"
                     /note="Inverse PCR was performed on Drosophila
                     melanogaster strains each of which contains one or more
                     P{GT1} P-element transposon insertion.  The resultant
                     fragment for each strain was directly sequenced to
                     determine the genomic sequence at the site of insertion.
                     Details of the protocols used can be found at
                     http://www.fruitfly.org/about/methods/inverse.pcr.html."
BASE COUNT       21 a     15 c      9 g     25 t
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 17;  Length 70;
  Best Local Similarity   83.3%;  Pred. No. 6.5e+03;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       3 GGACAAATCGAGTTATCA 20
             | ||||||| |||| |||
Db      22 GCACAAATCCAGTTCTCA 39


RESULT 6
AI222966
LOCUS       AI222966                 76 bp    mRNA    linear   EST 30-NOV-1998
DEFINITION  qh23e04.x1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone
            IMAGE:1845534 3' similar to TR:Q15941 Q15941 ZONA-PELLUCIDA-BINDING
            PROTEIN ;, mRNA sequence.
ACCESSION   AI222966
VERSION     AI222966.1  GI:3805169
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 76)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Trace considered overall poor quality
            Insert Length: 742   Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .76
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1845534"
                     /clone_lib="Soares_NFL_T_GBC_S1"
                     /lab_host="DH10B"
                     /note="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with
                     a modified polylinker; Site_1: Not I; Site_2: Eco RI;
                     Equal amounts of plasmid DNA from three normalized
                     libraries (fetal lung NbHL19W, testis NHT, and B-cell
                     NCI_CGAP_GCB1) were mixed, and ss circles were made in
```

vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 3 libraries. The pools consisted of I.M.A.G.E. clones 297480-302087, 682632-687239, 726408-728711, and 729096-731399. Subtraction by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT 23 a 15 c 18 g 20 t
ORIGIN

Query Match 66.0%; Score 13.2; DB 9; Length 76;
Best Local Similarity 83.3%; Pred. No. 6.8e+03;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3 GGACAAATCGAGTTATCA 20
| ||||||| || |||||
Db 28 GAACAAATCTAGGTATCA 45

RESULT 7
AA509111/c
LOCUS AA509111 100 bp mRNA linear EST 08-JUL-1997
DEFINITION MBAFCX7D08T3 Brugia malayi adult female cDNA (SAW96MLW-BmAF) Brugia malayi cDNA clone AFCX7D08 5', mRNA sequence.
ACCESSION AA509111
VERSION AA509111.1 GI:2246988
KEYWORDS EST.
SOURCE Brugia malayi.
ORGANISM Brugia malayi
Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea; Onchocercidae; Brugia.
REFERENCE 1 (bases 1 to 100)
AUTHORS Blaxter,M.L., Waterfall,M., Daub,J., Lizotte,M., Baron,L. and Jones,S.J.
TITLE Genes expressed in adult female Brugia malayi
JOURNAL Unpublished (1996)
COMMENT Contact: Blaxter ML
Institute of Cell, Animal and Population Biology
University of Edinburgh
Ashworth Labs, King's Buildings, West Mains Road, Edinburgh, EH9 3JT, UK.
Tel: +44 131 650 6760
Fax: +44 131 670 5450
Email: mark.blaxter@ed.ac.uk
The ABI trace of this sequence can be viewed at http://www.sanger.ac.uk/brugia/AFC/MBAFCX7D08T3.html
Seq primer: T3.
FEATURES Location/Qualifiers
source 1. .100
/organism="Brugia malayi"
/db_xref="taxon:6279"
/clone="AFCX7D08"
/clone_lib="Brugia malayi adult female cDNA (SAW96MLW-BmAF)"
/sex="female"
/dev_stage="adult"
/lab_host="XL1-Blue MRF'"
/note="Vector: Lambda Uni-ZAP XR; Site_1: Eco RI; Site_2: Xho I; Lymphatic filarial nematode parasite of humans. mRNA was prepared from approximately 50 adult females isolated from the peritoneal cavity of jirds and converted to double-stranded cDNA using reverse transcriptase and oligo(dT) followed by RNase H and DNA pol I. The library has 5 x 10E6 independent recombinants and the average insert size is ~900bp.The library was constructed by Michelle Lizotte-Waniewski. The library is available from Dr.S.A.Williams, email: genome@smith.edu."
BASE COUNT 27 a 17 c 15 g 41 t
ORIGIN

Query Match 66.0%; Score 13.2; DB 9; Length 100;
Best Local Similarity 83.3%; Pred. No. 7.6e+03;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3 GGACAAATCGAGTTATCA 20
||||||| ||| |||||
Db 56 GGACAAAGTGAGATATCA 39

RESULT 8
BH812020
LOCUS BH812020 79 bp DNA linear GSS 02-MAY-2002
DEFINITION SALK_061003 Arabidopsis thaliana TDNA insertion lines Arabidopsis thaliana genomic clone SALK_061003, DNA sequence.
ACCESSION BH812020
VERSION BH812020.1 GI:20390475
KEYWORDS GSS.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 79)
AUTHORS Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R., Gadrinab,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L., Shinn,P., Zimmerman,J. and Ecker,J.R.
TITLE A Sequence-Indexed Library of Insertion Mutations in the Arabidopsis Genome
JOURNAL Unpublished (2001)
COMMENT Contact: Joseph R. Ecker
Salk Institute Genomic Analysis Laboratory (SIGnAL)
The Salk Institute for Biological Studies
10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
Tel: 858 453 4100 x1752
Fax: 858 558 6379
Email: ecker@salk.edu
This is single pass sequence recovered from the left border of TDNA. This sequence lies within an annotated exon of At4g21370.
Class: TDNA tagged.
FEATURES Location/Qualifiers
source 1. .79
/organism="Arabidopsis thaliana"
/strain="Columbia 0"
/db_xref="taxon:3702"
/clone="SALK_061003"
/clone_lib="Arabidopsis thaliana TDNA insertion lines"
/note="PCR was performed on Arabidopsis thaliana lines each of which contains one or more TDNA insertion elements. The resultant fragment for each line was directly sequenced to determine the genomic sequence at the site of insertion. Details of the protocols used can be found at http://signal.salk.edu/tdna_protocols.html"
BASE COUNT 29 a 14 c 15 g 21 t
ORIGIN

Query Match 64.0%; Score 12.8; DB 17; Length 79;
Best Local Similarity 87.5%; Pred. No. 1.1e+04;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 5 ACAAATCGAGTTATCA 20
|||||| | |||||||
Db 28 ACAAATTGTGTTATCA 43

RESULT 9
AZ661477
LOCUS AZ661477 62 bp DNA linear GSS 14-DEC-2000
DEFINITION 1M0540A14F Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC1M0540A14 F, DNA sequence.
ACCESSION AZ661477
VERSION AZ661477.1 GI:11798719
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 62)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0540  row: A  column: 14
            Seq primer: CGTTGTAAAACGACGGCCAGT
            Class: plasmid ends
            High quality sequence stop: 62.
FEATURES             Location/Qualifiers
     source          1. .62
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC1M0540A14"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
BASE COUNT       26 a      7 c     11 g     18 t
ORIGIN

  Query Match             63.0%;  Score 12.6;  DB 17;  Length 62;
  Best Local Similarity   78.9%;  Pred. No. 1.3e+04;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 TCGGACAAATCGAGTTATC 19
           | |||||||| ||  ||||
Db      30 TTGGACAAATGGATGTATC 48


RESULT 10
AU031313/c
LOCUS       AU031313                 75 bp    mRNA    linear   EST 01-APR-2002
DEFINITION  AU031313 Rice cDNA from immature leaf including apical meristem
            Oryza sativa (japonica cultivar-group) cDNA clone E61323_2Z, mRNA
            sequence.
ACCESSION   AU031313
VERSION     AU031313.1  GI:3767203
KEYWORDS    EST.
SOURCE      Oryza sativa (japonica cultivar-group).
  ORGANISM  Oryza sativa (japonica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1  (bases 1 to 75)
  AUTHORS   Sasaki,T. and Yamamoto,K.
  TITLE     Rice cDNA from immature leaf including apical meristem
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Takuji Sasaki
            National Institute of Agrobiological Resources
            Rice Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki
            305-8602, Japan
            Tel: 81-298-38-7441
            Fax: 81-298-38-7468
            Email: tsasaki@abr.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/
            PROJECT ='RGP'
            POLYA=No.
FEATURES             Location/Qualifiers
     source          1. .75
                     /organism="Oryza sativa (japonica cultivar-group)"
                     /cultivar="Nipponbare"
                     /db_xref="taxon:39947"
                     /clone="E61323_2Z"
                     /clone_lib="Rice cDNA from immature leaf including apical
                     meristem"
                     /dev_stage="immature"
                     /note="Organ: leaf; immature leaf including apical
                     meristem (under long day condition)"
BASE COUNT       31 a     10 c     10 g     23 t      1 others
ORIGIN

  Query Match             63.0%;  Score 12.6;  DB 9;  Length 75;
  Best Local Similarity   75.0%;  Pred. No. 1.4e+04;
  Matches   15;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1 TCGGACAAATCGAGTTATCA 20
           | ||| |||| ||  |||||
Db      62 TTGGANAAATTGATGTATCA 43


RESULT 11
AL757650/c
LOCUS       AL757650                 97 bp    DNA     linear   GSS 18-JUN-2002
DEFINITION  Arabidopsis thaliana T-DNA flanking sequence GK-151C10-013133,
            genomic survey sequence.
ACCESSION   AL757650
VERSION     AL757650.1  GI:21495998
KEYWORDS    GSS.
SOURCE      thale cress.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1
  AUTHORS   Strizhov,N., Li,Y., Rosso,M., Viehoever,P., Dekker,K., Saedler,H.
            and Weisshaar,B.
  TITLE     A pipeline for automated high-throughput generation of FSTs
            (flanking sequence tags) from Arabidopsis thaliana T-DNA
            transformed lines
  JOURNAL   Unpublished
REFERENCE   2
  AUTHORS   Rosso,M., Strizhov,N., Li,Y., Reiss,B., Dekker,K. and Weisshaar,B.
  TITLE     A new Arabidopsis thaliana T-DNA mutagenised population (GABI-Kat)
            for flanking sequence tag based reverse genetics
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 97)
  AUTHORS   Li,Y., Rosso,M., Strizhov,N. and Weisshaar,B.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JUN-2002) Weisshaar B., Max-Planck-Institut fuer
            Zuechtungsforschung, Carl-von-Linne-Weg 10, Koeln, 50829, Germany
COMMENT     This sequence is recovered from the left border of the T-DNA. It
            indicates an insertion close to or within gene At5g43170. The
            sequences are generated at the MPI for Plant Breeding Research in
            the context of the GABI-Kat project. GABI-Kat is part of the German
```

```
            Plant Genomics program designated 'GABI'. Information on line
            availability can be found at:
            http://www.mpiz-koeln.mpg.de/GABI-Kat/.
FEATURES              Location/Qualifiers
     source           1. .97
                      /organism="Arabidopsis thaliana"
                      /strain="Columbia 0"
                      /db_xref="taxon:3702"
                      /clone="GK-151C10-013133"
                      /clone_lib="Arabidopsis thaliana T-DNA insertion lines"
                      /note="PCR was performed on DNA from Arabidopsis thaliana
                      plants (T1) which were transformed with the T-DNA from
                      vector pAC161. The lines contain one or more T-DNA
                      insertions. The DNA fragment(s) resulting from the PCR
                      were directly sequenced to determine the genomic sequence
                      flanking the insertion. Sequences displaying significant
                      similarity to the A. thalaina nuclear genome sequence were
                      processed for submission. T-DNA derived sequences were
                      removed"
BASE COUNT       30 a     30 c     23 g     14 t
ORIGIN

  Query Match             63.0%;  Score 12.6;  DB 17;  Length 97;
  Best Local Similarity   78.9%;  Pred. No. 1.6e+04;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 CGGACAAATCGAGTTATCA 20
           |||| ||  | ||||||||
Db      29 CGGAGAAGCCCAGTTATCA 11


RESULT 12
AA671696/c
LOCUS       AA671696                100 bp    mRNA    linear   EST 25-NOV-1997
DEFINITION  vl09b07.r1 Soares_mammary_gland_NbMMG Mus musculus cDNA clone
            IMAGE:963637 5' similar to TR:E221193 E221193 CYTOCHROME C OXIDASE
            POLYPEPTIDE II ;, mRNA sequence.
ACCESSION   AA671696
VERSION     AA671696.1  GI:2643775
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 100)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:552429
            Trace considered overall poor quality
            Possible reversed clone: similarity on wrong strand
            Seq primer: -28m13 rev2 ET from Amersham
            High quality sequence stop: 1.
FEATURES              Location/Qualifiers
     source           1. .100
                      /organism="Mus musculus"
                      /strain="C57BL/6J"
                      /db_xref="taxon:10090"
                      /clone="IMAGE:963637"
                      /clone_lib="Soares_mammary_gland_NbMMG"
                      /sex="male"
                      /tissue_type="mammary gland"
                      /dev_stage="4 weeks"
                      /lab_host="DH10B"
                      /note="Organ: mammary gland; Vector: pT7T3D-Pac (Pharmacia
                      ) with a modified polylinker; Site_1: Not I; Site_2: Eco
                      RI; 1st strand cDNA was primed with a Not I - oligo(dT)
                      primer [5'
                      TGTTACCAATCTGAAGTGGGAGCGGCCGCGAATGGTTTTTTTTTTTTTTTTTTTT
                      T 3']; double-stranded cDNA was ligated to Eco RI
                      adaptors (Pharmacia), digested with Not I and cloned into
                      the Not I and Eco RI sites of the modified pT7T3 vector.
                      RNA provided by Dr. Minoru Ko, Wayne State Univ. Library
                      constructed and normalized by Bento Soares and M.Fatima
                      Bonaldo."
BASE COUNT       31 a     18 c     27 g     24 t
ORIGIN

  Query Match             63.0%;  Score 12.6;  DB 9;  Length 100;
  Best Local Similarity   78.9%;  Pred. No. 1.6e+04;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 TCGGACAAATCGAGTTATC 19
           | | |||| | ||||||||
Db      80 TTGTACAACTGGAGTTATC 62


RESULT 13
TA83D07Q
LOCUS       TA83D07Q                55 bp    DNA     linear   GSS 13-DEC-2000
DEFINITION  T. brucei sheared genomic DNA clone 83d07, reverse sequence,
            genomic survey sequence.
ACCESSION   AL462388
VERSION     AL462388.1  GI:11861023
KEYWORDS    GSS.
SOURCE      Trypanosoma brucei.
  ORGANISM  Trypanosoma brucei
            Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
            Trypanosoma.
REFERENCE   1  (bases 1 to 55)
  AUTHORS   Hall,N., Bowman,S., Lennard,N.J., Doggett,J., Atkin,R.,
            Chillingworth,C., Ormond,D., Harris,B., El-Sayed,N., Hou,L.,
            Melville,S.E., Rajandream,M.A. and Barrell,B.G.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-DEC-2000) Trypanosoma brucei genome sequencing
            project, Sanger Centre, The Wellcome Trust Genome Campus, Hinxton,
            Cambridge CB10 1SA, E-mail: barrell@sanger.ac.uk and
            nh1@sanger.ac.uk
COMMENT     Constructed at the Institute for Genomic Research (TIGR),
            Rockville, MD. Genomic DNA isolated from a cloned population of
            Trypanosoma brucei (TREU927/4 GUTat 10.1) was mechanically sheared
            to give a tight size distribution (
            4 kb). The v + i method used for the library construction is
            described in detail in Smith, H. and Venter, J.C. (Making small
            insert libraries for whole genome shotgun sequencing projects. In
            Genome Sequencing: A Practical Approach, eds. M. Vaudin and B.
            Barrell, Oxford University Press, 1999).
            Email: nelsayed@tigr.org
            Details of T. brucei sequencing at the Sanger Centre are available
            at http://www.sanger.ac.uk/Projects/T_brucei/.
FEATURES              Location/Qualifiers
     source           1. .55
                      /organism="Trypanosoma brucei"
                      /strain="TREU927"
                      /db_xref="taxon:5691"
                      /clone="83d07"
BASE COUNT       19 a     11 c     15 g     10 t
ORIGIN

  Query Match             62.0%;  Score 12.4;  DB 17;  Length 55;
  Best Local Similarity   92.9%;  Pred. No. 1.6e+04;
  Matches   13;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;
```

```
Qy       7 AAATCGAGTTATCA 20
           |||||||||| ||
Db       6 AAATCGAGTTACCA 19
```

```
RESULT 14
AV832570/c
LOCUS       AV832570                  57 bp    mRNA    linear   EST 09-MAY-2002
DEFINITION  AV832570 K. Sato unpublished cDNA library: Hordeum vulgare subsp.
            vulgare leaves vegetative stage Hordeum vulgare subsp. vulgare cDNA
            clone baak13n05, mRNA sequence.
ACCESSION   AV832570
VERSION     AV832570.1  GI:14524659
KEYWORDS    EST.
SOURCE      Hordeum vulgare subsp. vulgare.
  ORGANISM  Hordeum vulgare subsp. vulgare
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
            ; Triticeae; Hordeum.
REFERENCE   1  (bases 1 to 57)
  AUTHORS   Sato,K.
  TITLE     Barley EST sequencing project in NIG and Okayama Univ
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Kazuhiro Sato
            Research Institute for Bioresources
            Okayama University, Barley Germplasm Center
            Chuo 2-20-1, Kurashiki, Okayama 710-0046, Japan
            Email: kazsato@rib.okayama-u.ac.jp,
            URL:http://www.rib.okayama-u.ac.jp/barley/
            database:http://www.shigen.nig.ac.jp/barley/Barley.html.
FEATURES             Location/Qualifiers
     source          1. .57
                     /organism="Hordeum vulgare subsp. vulgare"
                     /cultivar="Akashinriki"
                     /db_xref="taxon:112509"
                     /clone="baak13n05"
                     /clone_lib="K. Sato unpublished cDNA library: Hordeum
                     vulgare subsp. vulgare leaves vegetative stage"
                     /tissue_type="leaves"
                     /dev_stage="vegetative stage"
BASE COUNT       16 a     14 c     13 g     14 t
ORIGIN
```

```
  Query Match           62.0%;  Score 12.4;  DB 10;  Length 57;
  Best Local Similarity  92.9%;  Pred. No. 1.6e+04;
  Matches   13;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       7 AAATCGAGTTATCA 20
           |||||||||| |||
Db      33 AAATCGAGTTCTCA 20
```

```
RESULT 15
AQ073044
LOCUS       AQ073044                  84 bp    DNA     linear   GSS 23-AUG-2000
DEFINITION  EP(3)3592 Drosophila melanogaster EP line Drosophila melanogaster
            genomic Sequence recovered from Both 5' and 3' ends of P element,
            DNA sequence.
ACCESSION   AQ073044
VERSION     AQ073044.1  GI:3404161
KEYWORDS    GSS.
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
            Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 84)
  AUTHORS   Liao,G.-C., Rehm,E.J. and Rubin,G.M.
  TITLE     Insertion site preferences of the P transposable element in
            Drosophila melanogaster
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3347-3351 (2000)
  MEDLINE   20202638
COMMENT     Contact: Gerald Rubin
            Berkeley Drosophila Genome Project
            University of California, Berkeley
            LSA Building, Berkeley, CA 94720-3200, USA
            Fax: 5106439947
            Email: gerry@fruitfly.berkeley.edu
            Sequence recovery method was inverse PCR.

            Sequence orientation is forward strand relative to 5' end of P
            element

            The P element insertion position is base 3 in the 84 bases. This
            insertion position refers to the first base of the 8 base target
            recognition sequence.
            Class: transposon-tagged.
FEATURES             Location/Qualifiers
     source          1. .84
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone_lib="Drosophila melanogaster EP line"
                     /note="Inverse PCR was performed on Drosophila
                     melanogaster strains each of which contains a single EP
                     transposable element insertion. (The generation of these
                     insertion strains is described in Rorth P, Szabo K, Bailey
                     A, Laverty T, Rehm J, Rubin GM, Weigmann K, Milan M, Benes
                     V, Ansorge W, Cohen SM. 1998. Systematic gain-of-function
                     genetics in Drosophila. Development 6:1049-1057.) The
                     resultant fragment for each strain was directly sequenced
                     to determine the genomic sequence at the site of
                     insertion. Details of the protocols used can be found at
                     http://fruitfly.berkeley.edu/p_disrupt/inverse_pcr.html."
BASE COUNT       34 a     19 c     23 g      8 t
ORIGIN
```

```
  Query Match           62.0%;  Score 12.4;  DB 17;  Length 84;
  Best Local Similarity  92.9%;  Pred. No. 1.9e+04;
  Matches   13;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 CGGACAAATCGAGT 15
           | ||||||||||||
Db      31 CAGACAAATCGAGT 44
```

```
Search completed: November 26, 2002, 17:57:31
Job time : 810.5 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 3, 2002, 18:14:22 ; Search time 351.3 Seconds
(without alignments)
1656.863 Million cell updates/sec

Title: US-09-296-264-24
Perfect score: 20
Sequence: 1 caacattccagagcaaggat 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 995600

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
c	1	14.2	71.0	65	6	AX485936	AX485936 Sequence
	2	14	70.0	84	9	HSZ92872	Z92872 Homo sapien
c	3	13.8	69.0	17	6	AX099957	AX099957 Sequence
	4	13.8	69.0	51	6	AX203958	AX203958 Sequence
c	5	13.6	68.0	31	6	AR090047	AR090047 Sequence
c	6	13.6	68.0	31	6	AR197082	AR197082 Sequence
c	7	13.6	68.0	36	6	E08109	E08109 Synthetic o
c	8	13.6	68.0	36	6	E09916	E09916 Primer for
	9	13.6	68.0	92	17	HSMC24E11	X88251 H.sapiens D
c	10	13.2	66.0	29	6	AX298101	AX298101 Sequence
c	11	13.2	66.0	51	6	AX163349	AX163349 Sequence
	12	13	65.0	88	11	HSC124EA	Z78191 H.sapiens D
	13	13	65.0	96	11	HSC124EH4	Z78193 H.sapiens D
c	14	12.8	64.0	18	6	AR112304	AR112304 Sequence
	15	12.8	64.0	20	6	AR153204	AR153204 Sequence
	16	12.8	64.0	20	6	AR163278	AR163278 Sequence
	17	12.8	64.0	20	6	AR176944	AR176944 Sequence
	18	12.8	64.0	21	6	AR193717	AR193717 Sequence
c	19	12.8	64.0	26	6	AR107908	AR107908 Sequence
	20	12.8	64.0	29	6	AX103768	AX103768 Sequence
	21	12.8	64.0	29	6	AX461495	AX461495 Sequence
	22	12.8	64.0	31	6	AX249615	AX249615 Sequence
c	23	12.8	64.0	40	6	AR009896	AR009896 Sequence
c	24	12.8	64.0	51	6	AX165681	AX165681 Sequence
c	25	12.8	64.0	67	6	A79761	A79761 Sequence 9
c	26	12.8	64.0	67	6	AR135774	AR135774 Sequence
	27	12.8	64.0	67	6	AR135775	AR135775 Sequence
	28	12.8	64.0	92	14	HTU12119	U12119 Human T-cel
	29	12.8	64.0	93	6	AX127784	AX127784 Sequence
	30	12.8	64.0	93	6	AX139771	AX139771 Sequence
	31	12.6	63.0	23	6	I95942	I95942 Sequence 17
	32	12.6	63.0	23	6	I95967	I95967 Sequence 42
c	33	12.6	63.0	23	6	I95993	I95993 Sequence 68
c	34	12.6	63.0	23	6	I96018	I96018 Sequence 93
c	35	12.6	63.0	24	6	AX036495	AX036495 Sequence
c	36	12.6	63.0	27	6	AR109697	AR109697 Sequence
	37	12.6	63.0	30	6	AR053305	AR053305 Sequence
	38	12.6	63.0	30	6	I40924	I40924 Sequence 39
c	39	12.6	63.0	40	12	SYN322SVC	K01421 pBR322/SV40
c	40	12.6	63.0	60	6	A29449	A29449 sequence ex
c	41	12.6	63.0	60	6	I03045	I03045 Sequence 12
c	42	12.6	63.0	78	6	I03046	I03046 Sequence 13
	43	12.6	63.0	83	6	AX092857	AX092857 Sequence
c	44	12.6	63.0	90	6	AX368578	AX368578 Sequence
	45	12.6	63.0	91	6	AX092858	AX092858 Sequence

ALIGNMENTS

RESULT 1
AX485936/c
LOCUS AX485936 65 bp DNA linear PAT 16-AUG-2002
DEFINITION Sequence 3236 from Patent WO02053728.
ACCESSION AX485936
VERSION AX485936.1 GI:22320152
KEYWORDS .
SOURCE Candida albicans.
ORGANISM Candida albicans
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; mitosporic Saccharomycetales; Candida.
REFERENCE 1
AUTHORS Roemer,T., Jiang,B., Boone,C., Bussey,H. and Ohlsen,K.L.
TITLE Gene disruption methodologies for drug target discovery
JOURNAL Patent: WO 02053728-A 3236 11-JUL-2002;

```
           Elitra Pharmaceuticals, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .65
                     /organism="Candida albicans"
                     /db_xref="taxon:5476"
BASE COUNT       19 a     12 c     10 g     24 t
ORIGIN

  Query Match            71.0%;  Score 14.2;  DB 6;  Length 65;
  Best Local Similarity   84.2%;  Pred. No. 7.9e+03;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 AACATTCCAGAGCAAGGAT 20
           |||| | |||||||||| ||
Db      50 AACACTGCAGAGCAAGCAT 32


RESULT 2
HSZ92872
LOCUS       HSZ92872                 84 bp    DNA     linear   PRI 16-APR-1998
DEFINITION  Homo sapiens PRKG1 gene, exon 5.
ACCESSION   Z92872
VERSION     Z92872.1  GI:3063846
KEYWORDS    cGMP-dependent protein kinase; PRKG1 gene.
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 84)
  AUTHORS   Orstavik,S., Natarajan,V., Tasken,K., Jahnsen,T. and Sandberg,M.
  TITLE     Characterization of the human gene encoding the type I alpha and
            type I beta cGMP-dependent protein kinase (PRKG1)
  JOURNAL   Genomics 42 (2), 311-318 (1997)
  MEDLINE   97336057
   PUBMED   9192852
REFERENCE   2  (bases 1 to 84)
  AUTHORS   Orstavik,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-FEB-1998) Orstavik S., Institute of Medical
            Biochemistry, University of Oslo, pb 1112, Blindern, N-0317 Oslo,
            NORWAY
COMMENT     Related sequence: Y07512.
FEATURES             Location/Qualifiers
     source          1. .84
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="10"
                     /map="q11.1"
                     /clone="lambda 146"
                     /clone_lib="lambda EMBL3"
     gene            1. .84
                     /gene="PRKG1"
     intron          <1. .10
                     /gene="PRKG1"
                     /number=4
     exon            11. .74
                     /gene="PRKG1"
                     /number=5
                     /usedin=Z92867:CGKIA_CDS
                     /usedin=Z92868:CGKIB_CDS
     intron          75. .>84
                     /gene="PRKG1"
                     /number=5
BASE COUNT       19 a     19 c     18 g     28 t
ORIGIN

  Query Match            70.0%;  Score 14;  DB 9;  Length 84;
  Best Local Similarity   100.0%;  Pred. No. 1e+04;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CAACATTCCAGAGC 14
           ||||||||||||||
Db      16 CAACATTCCAGAGC 29


RESULT 3
AX099957/c
LOCUS       AX099957                 17 bp    DNA     linear   PAT 02-APR-2001
DEFINITION  Sequence 17 from Patent WO0120034.
ACCESSION   AX099957
VERSION     AX099957.1  GI:13538967
KEYWORDS    .
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 17)
  AUTHORS   Voss,J. and Timm,J.
  TITLE     Methods and compositions for the screening of cell cycle modulators
  JOURNAL   Patent: WO 0120034-A 17 22-MAR-2001;
            BASF AKTIENGESELLSCHAFT (DE)
FEATURES             Location/Qualifiers
     source          1. .17
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
BASE COUNT        4 a      3 c      4 g      6 t
ORIGIN

  Query Match            69.0%;  Score 13.8;  DB 6;  Length 17;
  Best Local Similarity   88.2%;  Pred. No. 1.2e+04;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 CAACATTCCAGAGCAAG 17
           || ||||| ||||||||
Db      17 CATCATTCTAGAGCAAG 1


RESULT 4
AX203958
LOCUS       AX203958                 51 bp    DNA     linear   PAT 30-AUG-2001
DEFINITION  Sequence 64 from Patent WO0148245.
ACCESSION   AX203958
VERSION     AX203958.1  GI:15393421
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 51)
  AUTHORS   Shimkets,R.A. and Leach,M.
  TITLE     Nucleic acids containing single nucleotide polymorphisms and
            methods of use thereof
  JOURNAL   Patent: WO 0148245-A 64 05-JUL-2001;
            Curagen Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .51
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     variation       26
                     /note="single nucleotide polymorphism
                     Accession number cg43090990"
BASE COUNT       13 a     11 c      8 g     19 t
ORIGIN

  Query Match            69.0%;  Score 13.8;  DB 6;  Length 51;
  Best Local Similarity   88.2%;  Pred. No. 1.3e+04;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       3 ACATTCCAGAGCAAGGA 19
           ||||| |||| ||||||
Db      21 ACATTTCAGACCAAGGA 37


RESULT 5
AR090047/c
LOCUS       AR090047                 31 bp    DNA     linear   PAT 07-SEP-2000
```

```
DEFINITION  Sequence 167 from patent US 5994076.
ACCESSION   AR090047
VERSION     AR090047.1  GI:10016802
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 31)
  AUTHORS   Chenchik,A., Jokhadze,G. and Bibilashvilli,R.
  TITLE     Methods of assaying differential expression
  JOURNAL   Patent: US 5994076-A 167 30-NOV-1999;
FEATURES             Location/Qualifiers
     source          1. .31
                     /organism="unknown"
BASE COUNT        5 a      8 c      7 g     11 t
ORIGIN

  Query Match            68.0%;  Score 13.6;  DB 6;  Length 31;
  Best Local Similarity  80.0%;  Pred. No. 1.7e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 CAACATTCCAGAGCAAGGAT 20
          || |  |||||||||| |||
Db     23 CAGCGGTCCAGAGCAATGAT 4


RESULT 6
AR197082/c
LOCUS       AR197082                  31 bp    DNA     linear   PAT 20-APR-2002
DEFINITION  Sequence 167 from patent US 6352829.
ACCESSION   AR197082
VERSION     AR197082.1  GI:20246931
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 31)
  AUTHORS   Chenchik,A., Jokhadze,G. and Bibilashvilli,R.
  TITLE     Methods of assaying differential expression
  JOURNAL   Patent: US 6352829-A 167 05-MAR-2002;
FEATURES             Location/Qualifiers
     source          1. .31
                     /organism="unknown"
BASE COUNT        5 a      8 c      7 g     11 t
ORIGIN

  Query Match            68.0%;  Score 13.6;  DB 6;  Length 31;
  Best Local Similarity  80.0%;  Pred. No. 1.7e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 CAACATTCCAGAGCAAGGAT 20
          || |  |||||||||| |||
Db     23 CAGCGGTCCAGAGCAATGAT 4


RESULT 7
E08109/c
LOCUS       E08109                    36 bp    DNA     linear   PAT 29-SEP-1997
DEFINITION  Synthetic oligonucleotides for probe.
ACCESSION   E08109
VERSION     E08109.1  GI:2176230
KEYWORDS    JP 1994253858-A/12.
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 36)
  AUTHORS   Beppu,T., Horinouchi,S. and Nomura,N.
  TITLE     FUSED GENE HAVING RODENTIAN APOLIPOPROTEIN E, PLASMID COMPRISING
            THE SAME INSERTED THEREINTO AND USE THEREOF
  JOURNAL   Patent: JP 1994253858-A 12 13-SEP-1994;
            BEPPU TERUHIKO
COMMENT     OS   None
            OC   Artificial sequences.
            PN   JP 1994253858-A/12
            PD   13-SEP-1994
            PF   04-MAR-1993 JP 1993069446
            PI   BEPPU TERUHIKO, HORINOUCHI SUEJI, NOMURA NOBUHIKO PC
            C12N15/62,C12N1/21,C12N9/64,C12N15/12,C12N15/57,C12P21/02; CC
            strandedness: Single;
            CC   topology: Linear;
            FH   Key             Location/Qualifiers
            FH
            FT   source          1. .36
            FT                   /organism='Artificial sequences'.
FEATURES             Location/Qualifiers
     source          1. .36
                     /organism="unidentified"
                     /db_xref="taxon:32644"
BASE COUNT        5 a      8 c     10 g     13 t
ORIGIN

  Query Match            68.0%;  Score 13.6;  DB 6;  Length 36;
  Best Local Similarity  80.0%;  Pred. No. 1.7e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 CAACATTCCAGAGCAAGGAT 20
          ||||| || |||  ||||||
Db     23 CAACAGTCTAGACTAAGGAT 4


RESULT 8
E09916/c
LOCUS       E09916                    36 bp    DNA     linear   PAT 29-SEP-1997
DEFINITION  Primer for gaining plasmid pHMAE.
ACCESSION   E09916
VERSION     E09916.1  GI:22026544
KEYWORDS    JP 1995241195-A/10.
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 36)
  AUTHORS   Beppu,T., Horinouchi,S. and Nomura,N.
  TITLE     FUSION GENE HAVING LYSOMUCORPEPSIN-LIKE GENE AND HUMAN
            APOLIPOPROTEIN GENE, PLASMID INCLUDING THE GENE AND ITS USE
  JOURNAL   Patent: JP 1995241195-A 10 19-SEP-1995;
            BEPPU TERUHIKO
COMMENT     OS   None
            OC   Artificial sequences.
            PN   JP 1995241195-A/10
            PD   19-SEP-1995
            PF   04-MAR-1994 JP 1994058269
            PI   BEPPU TERUHIKO, HORINOUCHI SUEJI, NOMURA NOBUHIKO PC
            C12N15/09,C12N1/19,C12N9/64,C12P21/02//A61K38/46,(C12N15/09, PC
            C12R1:645),
            PC   (C12N1/19,C12R1:645),(C12N9/64,C12R1:645),(C12P21/02, PC
            C12R1:645);
            CC   strandedness: Single;
            CC   topology: Linear;
            FH   Key             Location/Qualifiers
            FH
            FT   source          1. .36
            FT                   /organism='Artificial sequences' FT
            misc_feature    2. .7
            FT                   /note='BamHI digestion site'
            FT   misc_feature    12. .17
            FT                   /note='XbaI digestion site'
            FT   misc_feature    1. .36
            FT                   /note='Primer named MPP3'.
FEATURES             Location/Qualifiers
     source          1. .36
                     /organism="unidentified"
                     /db_xref="taxon:32644"
BASE COUNT        5 a      8 c     10 g     13 t
ORIGIN

  Query Match            68.0%;  Score 13.6;  DB 6;  Length 36;
```

```
  Best Local Similarity   80.0%;  Pred. No. 1.7e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 CAACATTCCAGAGCAAGGAT 20
           ||||| || |||  ||||||
Db      23 CAACAGTCTAGACTAAGGAT 4


RESULT 9
HSMC24E11
ID   HSMC24E11  standard; DNA; HUM; 92 BP.
XX
AC   X88251;
XX
SV   X88251.1
XX
DT   18-JUL-1996 (Rel. 48, Created)
DT   29-MAY-1997 (Rel. 52, Last updated, Version 3)
XX
DE   H.sapiens DNA for trapped exon (ID HMC24E11)
XX
KW   trapped exon.
XX
OS   Homo sapiens (human)
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia;
OC   Eutheria; Primates; Catarrhini; Hominidae; Homo.
XX
RN   [1]
RP   1-92
RA   Chen H.M., Rossier C., Chrast R., Antonarakis S.E.;
RT   "Cloning of trapped exons from human chromosome 21";
RL   Unpublished.
XX
RN   [2]
RP   1-92
RA   Antonarakis S.E.;
RT   ;
RL   Submitted (17-MAR-1995) to the EMBL/GenBank/DDBJ databases.
RL   Stylianos E. Antonarakis, Division of Medical Genetics, University and
RL   Cantonal Hospital of Geneva, CMU, 1 rue Michel-Servet, 1211 Geneva,
RL   SWITZERLAND
XX
RN   [3]
RP   1-92
RA   Chen H.M.; Rossier C., Chrast R., Antonarakis S.E.;
RT   "Cloning of 559 potential exons of genes of human chromosome 21 by exon
RT   trapping.";
RL   Genome Res. 6:747-760(1996).
XX
FH   Key             Location/Qualifiers
FH
FT   source          1. .92
FT                   /chromosome="21"
FT                   /db_xref="taxon:9606"
FT                   /organism="Homo sapiens"
FT   exon            1. .92
FT                   /note="trapped exon"
XX
SQ   Sequence 92 BP; 26 A; 30 C; 21 G; 14 T; 1 other;

  Query Match            68.0%;  Score 13.6;  DB 17;  Length 92;
  Best Local Similarity  80.0%;  Pred. No. 1.7e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 CAACATTCCAGAGCAAGGAT 20
           |||| | |||| || |||||
Db      48 CAACCTGCCAGTGCCAGGAT 67


RESULT 10
AX298101/c
LOCUS       AX298101                 29 bp    DNA     linear   PAT 26-NOV-2001
DEFINITION  Sequence 14 from Patent WO0183792.
ACCESSION   AX298101
VERSION     AX298101.1  GI:17128177
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Huang,N., Hwang,Y.S., Yang,D. and Schmidt,R.J.
  TITLE     Plant transcription factors and enhanced gene expression
  JOURNAL   Patent: WO 0183792-A 14 08-NOV-2001;
            Applied Phytologics, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .29
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="primer"
BASE COUNT        9 a      6 c      5 g      9 t
ORIGIN

  Query Match            66.0%;  Score 13.2;  DB 6;  Length 29;
  Best Local Similarity  83.3%;  Pred. No. 2.8e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 CAACATTCCAGAGCAAGG 18
           || || || |||||||||
Db      19 CATCAATCTAGAGCAAGG 2


RESULT 11
AX163349/c
LOCUS       AX163349                 51 bp    DNA     linear   PAT 22-JUN-2001
DEFINITION  Sequence 6677 from Patent WO0140521.
ACCESSION   AX163349
VERSION     AX163349.1  GI:14544680
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 51)
  AUTHORS   Shimkets,R.A. and Leach,M.
  TITLE     Nucleic acids containing single nucleotide polymorphisms and
            methods of use thereof
  JOURNAL   Patent: WO 0140521-A 6677 07-JUN-2001;
            Curagen Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .51
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     misc_feature    26
                     /note="1 of 2 allelic variants (6678 is other entry)
                     Accession number cg43103077"
BASE COUNT       10 a     10 c     12 g     19 t
ORIGIN

  Query Match            66.0%;  Score 13.2;  DB 6;  Length 51;
  Best Local Similarity  83.3%;  Pred. No. 2.9e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 AACATTCCAGAGCAAGGA 19
           | |||||||||| ||| ||
Db      34 AGCATTCCAGACCAAAGA 17


RESULT 12
HSC124EA
LOCUS       HSC124EA                 88 bp    DNA     linear   STS 13-AUG-1996
DEFINITION  H.sapiens DNA for STS c12/4EA, sequence tagged site.
ACCESSION   Z78191
VERSION     Z78191.1  GI:1490278
KEYWORDS    STS.
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
```

```
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 88)
  AUTHORS   Kostrzewa,M., Zakel,S. and Mueller,U.
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 88)
  AUTHORS   Kostrzewa,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-AUG-1996) Kostrzewa M., Institute of Human Genetics,
            Schlangenzahl 14, 35392 Giessen, FRG
FEATURES             Location/Qualifiers
     source          1. .88
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="5"
                     /map="5q34"
                     /clone="YAC 892A9"
                     /clone_lib="CEPH Mega YAC"
     STS             1. .88
                     /note="c12/4EA"
BASE COUNT       23 a     20 c     16 g     29 t
ORIGIN

  Query Match             65.0%;  Score 13;  DB 11;  Length 88;
  Best Local Similarity   100.0%;  Pred. No. 3.8e+04;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      4 CATTCCAGAGCAA 16
          |||||||||||||
Db     73 CATTCCAGAGCAA 85


RESULT 13
HSC124EH4
LOCUS       HSC124EH4                 96 bp    DNA     linear   STS 13-AUG-1996
DEFINITION  H.sapiens DNA for STS c12/4EH4, sequence tagged site.
ACCESSION   Z78193
VERSION     Z78193.1  GI:1490280
KEYWORDS    STS.
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 96)
  AUTHORS   Kostrzewa,M., Zakel,S. and Mueller,U.
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 96)
  AUTHORS   Kostrzewa,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-AUG-1996) Kostrzewa M., Institute of Human Genetics,
            Schlangenzahl 14, 35392 Giessen, FRG
FEATURES             Location/Qualifiers
     source          1. .96
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="5"
                     /map="5q34"
                     /clone="YAC 892A9"
                     /clone_lib="CEPH Mega YAC"
     STS             1. .96
                     /note="c12/4EH4"
BASE COUNT       23 a     24 c     18 g     31 t
ORIGIN

  Query Match             65.0%;  Score 13;  DB 11;  Length 96;
  Best Local Similarity   100.0%;  Pred. No. 3.8e+04;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      4 CATTCCAGAGCAA 16
          |||||||||||||
Db     75 CATTCCAGAGCAA 87


RESULT 14
AR112304/c
LOCUS       AR112304                  18 bp    DNA     linear   PAT 16-MAY-2001
DEFINITION  Sequence 25 from patent US 6130043.
ACCESSION   AR112304
VERSION     AR112304.1  GI:14092204
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 18)
  AUTHORS   Billing-Medel,P.A., Cohen,M., Colpitts,T.L., Friedman,P.N.,
            Gordon,J., Granados,E.N., Hodges,S.C., Klass,M.R., Kratochvil,J.D.,
            Roberts-Rapp,L., Russell,J.C., Stroupe,S.D. and Yu,H.
  TITLE     Reagents and methods useful for detecting diseases of the prostate
  JOURNAL   Patent: US 6130043-A 25 10-OCT-2000;
FEATURES             Location/Qualifiers
     source          1. .18
                     /organism="unknown"
BASE COUNT       4 a      5 c      5 g      4 t
ORIGIN

  Query Match             64.0%;  Score 12.8;  DB 6;  Length 18;
  Best Local Similarity   87.5%;  Pred. No. 4.6e+04;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      4 CATTCCAGAGCAAGGA 19
          |||||||| ||| |||
Db     18 CATTCCAGTGCATGGA 3


RESULT 15
AR153204
LOCUS       AR153204                  20 bp    DNA     linear   PAT 08-AUG-2001
DEFINITION  Sequence 206 from patent US 6235480.
ACCESSION   AR153204
VERSION     AR153204.1  GI:15120736
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 20)
  AUTHORS   Shultz,J.William., Lewis,M.K., Leippe,D., Mandrekar,M., Kephart,D.,
            Rhodes,R.Byron., Andrews,C.Ann., Hartnett,J.Robert., Gu,T.,
            Olson,R.J., Wood,K.V. and Welch,R.
  TITLE     Detection of nucleic acid hybrids
  JOURNAL   Patent: US 6235480-A 206 22-MAY-2001;
FEATURES             Location/Qualifiers
     source          1. .20
                     /organism="unknown"
BASE COUNT       5 a      4 c      7 g      4 t
ORIGIN

  Query Match             64.0%;  Score 12.8;  DB 6;  Length 20;
  Best Local Similarity   87.5%;  Pred. No. 4.6e+04;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      4 CATTCCAGAGCAAGGA 19
          | ||||||||| |||
Db      1 CTTTCCAGAGCAGGGA 16


Search completed: December  3, 2002, 22:23:20
Job time : 358.3 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 06:29:45 ; Search time 94.1 Seconds
(without alignments)
478.639 Million cell updates/sec

Title: US-09-296-264-24
Perfect score: 20
Sequence: 1 caacattccagagcaaggat 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 2390332

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	20	100.0	20	21	AAZ31454	Human neuropilin m
c	2	15	75.0	60	24	ABN48461	Human spliced tran
	3	14.2	71.0	41	22	AAH79193	Human interleukin
c	4	13.8	69.0	17	22	AAF57369	Murine Cdc25A intr
	5	13.8	69.0	51	22	AAH79449	Human DNA containi
c	6	13.8	69.0	60	24	ABN33165	Human spliced tran
c	7	13.8	69.0	60	24	ABN49462	Human spliced tran
c	8	13.6	68.0	31	24	ABK66079	Human gene specifi
c	9	13.6	68.0	36	15	AAQ72778	Primer MPR3 for am
c	10	13.6	68.0	36	16	AAT09079	Mucor pusillus pep
c	11	13.6	68.0	37	20	AAZ28379	PCR primer EGFPs u
c	12	13.6	68.0	48	20	AAX26807	Primer for 583/4/5
c	13	13.6	68.0	60	24	ABN36579	Human spliced tran
	14	13.6	68.0	84	22	ABA48643	Human breast cell
	15	13.6	68.0	84	22	ABA66556	Human foetal liver
	16	13.6	68.0	84	22	ABA33619	Probe #12085 for g
	17	13.6	68.0	84	22	AAK14979	Human brain expres
	18	13.6	68.0	84	22	AAK40714	Human bone marrow
	19	13.6	68.0	84	22	AAI21474	Probe #11407 for g
	20	13.6	68.0	84	22	AAI46765	Probe #15451 used
	21	13.6	68.0	84	22	AAI07172	Probe #7163 used t
	22	13.6	68.0	84	24	ABS14676	Human genome-deriv
	23	13.4	67.0	60	24	ABN48413	Human spliced tran
c	24	13.4	67.0	62	18	AAT60759	Bases (-225) to (-
c	25	13.2	66.0	29	24	ABA05214	OSglb PCR primer S
	26	13.2	66.0	31	22	AAI31206	Human single nucle
c	27	13.2	66.0	50	22	AAL29673	Human SNP oligonuc
	28	13.2	66.0	51	22	AAL28719	Human SNP oligonuc
c	29	13.2	66.0	51	22	AAI79736	Human nonconservat
	30	13.2	66.0	65	24	ABN30672	Rat spliced transc
c	31	13.2	66.0	96	19	AAV73226	C. utilis crtb pri
c	32	13.2	66.0	98	16	AAQ81660	bFGF binding oligo
c	33	12.8	64.0	18	21	AAC64937	Human prostate-rel
c	34	12.8	64.0	18	22	AAS07164	Prostate-specific
	35	12.8	64.0	20	21	AAA86995	CAH reverse probe
	36	12.8	64.0	20	21	AAA93320	Human congenital a
	37	12.8	64.0	20	24	AAD23033	Steroid 21-hydroxy
	38	12.8	64.0	21	14	AAQ43258	Sequence encoding
	39	12.8	64.0	21	14	AAQ43257	Sequence encoding
	40	12.8	64.0	21	15	AAQ71457	Rx 2.4 prorelaxin
c	41	12.8	64.0	26	17	AAT30199	3'-Primer for huma
c	42	12.8	64.0	28	21	AAA08335	Spinach chloroplas
	43	12.8	64.0	29	22	AAD03410	Wheat glutenin PrH
	44	12.8	64.0	33	22	AAI65169	Human retinaldehyd
	45	12.8	64.0	39	16	AAQ87104	Aspartic acid beta

ALIGNMENTS

RESULT 1
AAZ31454
ID AAZ31454 standard; DNA; 20 BP.
XX
AC AAZ31454;
XX
DT 07-FEB-2000 (first entry)
XX
DE Human neuropilin mRNA specific antisense oligo GTI3625.
XX
KW Neuropilin; human; growth; metastasis; tumor; neovascularisation;
KW cancer; papilloma; diabetic retinopathy; antisense; ss.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9955855-A2.
XX
PD 04-NOV-1999.
XX
PF 23-APR-1999; 99WO-CA00324.
XX
PR 23-APR-1998; 98US-0082791.
XX
PA (GENE-) GENESENSE TECHNOLOGIES INC.
XX
PI Wright JA, Young AH, Lee YS;
XX
DR WPI; 2000-023357/02.
XX
PT Antisense oligonucleotides that inhibit neuropilin expression, useful
PT for treating cancer -

XX
PS Claim 4; Page 17; 57pp; English.
XX
CC Sequences AAZ31431-460 represent antisense oligonucleotides which
CC inhibit human neuropilin expression. The antisense oligonucleotides can
CC be used to inhibit the growth or metastasis of a mammalian tumor and
CC inhibit neovascularisation. The oligonucleotides may be used to treat
CC various forms of cancers or tumors, such as sarcomas, melanomas,
CC adenomas, carcinomas of solid tissue, hypoxic tumors, squamous cell
CC carcinomas of the mouth, throat, larynx and lung, genitourinary cancers
CC such as cervical and bladder cancer, hematopoietic cancers, colon cancer,
CC breast cancer, pancreatic cancer, renal cancer, brain cancer, skin
CC cancer, liver cancer, head and neck cancers, and nervous system cancers,
CC as well as benign lesions such as papillomas. The methods may be used to
CC treat neovascularisation disorders such as diabetic retinopathy, and
CC retinopathy of prematurity and age related macular degeneration.
XX
SQ Sequence 20 BP; 8 A; 5 C; 4 G; 3 T; 0 other;

Query Match 100.0%; Score 20; DB 21; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.8;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CAACATTCCAGAGCAAGGAT 20
||||||||||||||||||||
Db 1 CAACATTCCAGAGCAAGGAT 20

RESULT 2
ABN48461/c
ID ABN48461 standard; DNA; 60 BP.
XX
AC ABN48461;
XX
DT 15-JUL-2002 (first entry)
XX
DE Human spliced transcript detection oligonucleotide SEQ ID NO:21209.
XX
KW Human; mouse; rat; splice transcript; detection; RNA transcript;
KW splice variant; transcriptome; oligonucleotide library; ss.
XX
OS Homo sapiens.
XX
PN WO200210449-A2.
XX
PD 07-FEB-2002.
XX
PF 20-JUL-2001; 2001WO-IB01903.
XX
PR 28-JUL-2000; 2000US-221607P.
PR 02-MAY-2001; 2001US-287724P.
XX
PA (COMP-) COMPUGEN INC.
XX
PI Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
XX
DR WPI; 2002-257383/30.
XX
PT New oligonucleotide libraries comprising oligonucleotides which
PT selectively hybridize to mRNAs transcribed from a transcription unit of
PT a genome, useful for detecting tissue-, pathology-, and
PT developmental-specific genes -
XX
PS Example 1; SEQ ID 21209; 47pp; English.
XX
CC The present invention describes oligonucleotide libraries for detecting
CC messenger RNAs that populate a (sub-)transcriptome, where the
CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC transcription units that populate a genome. The library comprises
CC several oligonucleotides, each capable of hybridising selectively to a
CC set of messenger RNAs transcribed from a given transcription unit of
CC the genome, which encodes one or more messenger RNA splice variants.
CC The oligonucleotide libraries are useful for detecting mRNAs from a
CC biological sample, in expression profiling studies, in qualitatively or
CC quantitatively characterising the corresponding transcriptome, and in
CC detecting RNA transcripts and splice variants of human or animal
CC transcriptomes. The libraries may also be used as specialised mini
CC libraries to detect transcripts of a sub-transcriptome under a
CC particular biological or pathological state, and so allowing the
CC detection of tissue- and pathology-specific genes such as those genes
CC only expressed in specific tissue under a specific pathological
CC condition; to detect developmental specific genes; and to detect RNA
CC transcripts and splice variants of a transcriptome of a patient suffering
CC from a particular disorder. ABN27253 to ABN59589 represent
CC oligonucleotide sequences from rats, humans and mice, which are used in
CC the exemplification of the present invention.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 60 BP; 15 A; 12 C; 13 G; 20 T; 0 other;

Query Match 75.0%; Score 15; DB 24; Length 60;
Best Local Similarity 100.0%; Pred. No. 6.1e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 AACATTCCAGAGCAA 16
|||||||||||||||
Db 58 AACATTCCAGAGCAA 44

RESULT 3
AAH79193
ID AAH79193 standard; DNA; 41 BP.
XX
AC AAH79193;
XX
DT 21-NOV-2001 (first entry)
XX
DE Human interleukin binding factor 1-12 probe 2.
XX
KW Human; interleukin binding factor 1-12; cytostatic; virucidal;
KW immunodulatory; antiinflammatory; haemostatic; malignant neoplasm;
KW HIV; human immunodeficiency virus; infection; immunological disease;
KW transfusion reaction; transplant immunological rejection;
KW major histocompatibility antigen-related disease; probe; ss.
XX
OS Homo sapiens.
XX
PN WO200166588-A1.
XX
PD 13-SEP-2001.
XX
PF 26-FEB-2001; 2001WO-CN00204.
XX
PR 07-MAR-2000; 2000CN-0111902.
XX
PA (BIOW-) BIOWINDOW GENE DEV INC SHANGHAI.
XX
PI Mao Y, Xie Y;
XX
DR WPI; 2001-596828/67.
XX
PT Human interleukin-binding factor 1-12 and encoded polynucleotide,
PT applicable in diagnosis and treatment of malignant neoplasm, haemopathy,
PT HIV infection, immunological diseases and various inflammations -
XX
PS Example 6; Page 15; 35pp; Chinese.
XX
CC The invention relates to human interleukin binding factor 1-12 and the
CC encoding polynucleotide with cytostatic, virucidal, immunodulatory,
CC antiinflammatory and haemostatic activity. The polypeptide and encoded
CC polynucleotide are applicable in diagnosis and treatment of malignant
CC neoplasm, haemopathy, HIV infection, immunological diseases and various
CC inflammations, particularly transfusion reaction, transplant
CC immunological rejection and major histocompatibility antigen-related

```
CC   diseases. The present sequence is that of a human interleukin binding
CC   factor 1-12 probe, useful to the invention.
XX
SQ   Sequence 41 BP; 15 A; 7 C; 14 G; 5 T; 0 other;

  Query Match             71.0%;  Score 14.2;  DB 22;  Length 41;
  Best Local Similarity   84.2%;  Pred. No. 1.5e+03;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 AACATTCCAGAGCAAGGAT 20
           ||||||  ||||||||| |
Db       9 AACATTGAAGAGCAAGGGT 27


RESULT 4
AAF57369/c
ID   AAF57369 standard; DNA; 17 BP.
XX
AC   AAF57369;
XX
DT   11-JUN-2001  (first entry)
XX
DE   Murine Cdc25A intron 7/exon 8 splice junction sequence.
XX
KW   Cdc25; Cdc25 phosphatase; transcription; modulator; murine; CDc25A;
KW   exon; intron; ds.
XX
OS   Mus sp.
XX
PN   WO200120034-A2.
XX
PD   22-MAR-2001.
XX
PF   11-SEP-2000; 2000WO-US24838.
XX
PR   13-SEP-1999;   99US-0153639.
XX
PA   (BADI ) BASF AG.
XX
PI   Voss J,  Timm J;
XX
DR   WPI; 2001-244825/25.
XX
PT   Assay for screening modulators of Cdc25 activity by using a cell having
PT   a recombinant Cdc25 phosphatase gene whose expression alters the
PT   transcription of a selected gene in the presence of a modulator -
XX
PS   Example 1; Page 15; 55pp; English.
XX
CC   The invention relates to a method of identifying a modulator of Cdc25
CC   activity that comprises contacting a test cell having a recombinant Cdc25
CC   phosphatase gene whose expression alters transcription of a selected
CC   gene, with a compound under conditions where recombinant Cdc25
CC   phosphatase gene is expressed and alters the transcription of a selected
CC   gene as an indication of the compound being a modulator of Cdc25-mediated
CC   transcription. The method is useful for identifying modulators of Cdc25
CC   activity. Sequences AAF57363-376 represent intron/exon splice junction
CC   sequences of the murine Cdc25A gene.
XX
SQ   Sequence 17 BP; 4 A; 3 C; 4 G; 6 T; 0 other;

  Query Match             69.0%;  Score 13.8;  DB 22;  Length 17;
  Best Local Similarity   88.2%;  Pred. No. 2.1e+03;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 CAACATTCCAGAGCAAG 17
           || ||||| ||||||||
Db      17 CATCATTCTAGAGCAAG 1


RESULT 5
AAH79449
ID   AAH79449 standard; DNA; 51 BP.
XX
AC   AAH79449;
XX
DT   19-SEP-2001  (first entry)
XX
DE   Human DNA containing single nucleotide polymorphism SEQ ID NO. 64.
XX
KW   Human; single nucleotide polymorphism; SNP; angiopoietin;
KW   4-hydroxybutyrate; dehydrogenase; protein therapy;
KW   adenosine triphosphate-dependent RNA helicase;
KW   major histocompatibility complex Class I histocompatibility antigen; MHC;
KW   phosphoglycerate kinase; immunosuppressive; immunostimulatory;
KW   antirheumatic; antisclerotic; antidiabetic; antiinflammatory; cytostatic;
KW   antileukemic; neuroprotective; antimicrobial; gene therapy; vaccine; ds.
XX
OS   Homo sapiens.
XX
PN   WO200148245-A2.
XX
PD   05-JUL-2001.
XX
PF   27-DEC-2000; 2000WO-US35346.
XX
PR   27-DEC-1999;   99US-0472688.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Shimkets RA,  Leach M;
XX
DR   WPI; 2001-418297/44.
XX
PT   Polymorphic nucleic acids encoding e.g. angiopoietin, dehydrogenase,
PT   adenosine triphosphate-dependent RNA helicase and/or phosphoglycerate
PT   kinase, useful for diagnosing and treating, e.g. cancer, autoimmune
PT   diseases and infections -
XX
PS   Claim 1; Page 76; 484pp; English.
XX
CC   The invention relates to nucleic acids (AAH79386-AAH80036) encoding
CC   polymorphic variants of proteins (AAG98010-AAG98238) related to
CC   angiopoietin, 4-hydroxybutyrate, dehydrogenase, adenosine triphosphate
CC   (ATP)-dependent RNA helicase, major histocompatibility complex (MHC)
CC   Class I histocompatibility antigen and/or phosphoglycerate kinase. These
CC   nucleic acid single nucleotide polymorphisms (SNPs) and the encoded
CC   proteins have potential immunosuppressive, immunostimulatory,
CC   antirheumatic, antisclerotic, antidiabetic, antiinflammatory, cytostatic,
CC   antileukemic, neuroprotective and antimicrobial activity and may be
CC   useful in gene/protein therapy, vaccines, modulation of the expression
CC   and activity of proteins related to angiopoietin, 4-hydroxybutyrate,
CC   dehydrogenase, adenosine triphosphate (ATP)-dependent RNA helicase,
CC   major histocompatibility complex (MHC) Class I histocompatibility antigen
CC   and/or phosphoglycerate kinase. Disorders that may be prevented,
CC   diagnosed and/or treated by the above methods include multifactorial
CC   diseases with a genetic component, such as autoimmune diseases (e.g.
CC   rheumatoid arthritis, multiple sclerosis, diabetes, systemic lupus
CC   erytheromatosus and Grave's disease), inflammation, cancer (e.g. cancers
CC   of the bladder, brain, breast, colon and kidney, leukemia), diseases of
CC   the nervous system, an infection of pathogenic organisms. They may also
CC   be used to alter phenotypic traits such as longevity, appearance,
CC   strength, speed and endurance.
XX
SQ   Sequence 51 BP; 13 A; 11 C; 8 G; 19 T; 0 other;

  Query Match             69.0%;  Score 13.8;  DB 22;  Length 51;
  Best Local Similarity   88.2%;  Pred. No. 2.4e+03;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       3 ACATTCCAGAGCAAGGA 19
           ||||| |||| ||||||
Db      21 ACATTTCAGACCAAGGA 37


RESULT 6
```

```
ABN33165/c
ID   ABN33165 standard; DNA; 60 BP.
XX
AC   ABN33165;
XX
DT   15-JUL-2002  (first entry)
XX
DE   Human spliced transcript detection oligonucleotide SEQ ID NO:5913.
XX
KW   Human; mouse; rat; splice transcript; detection; RNA transcript;
KW   splice variant; transcriptome; oligonucleotide library; ss.
XX
OS   Homo sapiens.
XX
PN   WO200210449-A2.
XX
PD   07-FEB-2002.
XX
PF   20-JUL-2001; 2001WO-IB01903.
XX
PR   28-JUL-2000; 2000US-221607P.
PR   02-MAY-2001; 2001US-287724P.
XX
PA   (COMP-) COMPUGEN INC.
XX
PI   Shoshan A,  Wasserman A,  Mintz E,  Mintz L,  Faigler S;
XX
DR   WPI; 2002-257383/30.
XX
PT   New oligonucleotide libraries comprising oligonucleotides which
PT   selectively hybridize to mRNAs transcribed from a transcription unit of
PT   a genome, useful for detecting tissue-, pathology-, and
PT   developmental-specific genes  -
XX
PS   Example 1; SEQ ID 5913; 47pp; English.
XX
CC   The present invention describes oligonucleotide libraries for detecting
CC   messenger RNAs that populate a (sub-)transcriptome, where the
CC   (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC   transcription units that populate a genome. The library comprises
CC   several oligonucleotides, each capable of hybridising selectively to a
CC   set of messenger RNAs transcribed from a given transcription unit of
CC   the genome, which encodes one or more messenger RNA splice variants.
CC   The oligonucleotide libraries are useful for detecting mRNAs from a
CC   biological sample, in expression profiling studies, in qualitatively or
CC   quantitatively characterising the corresponding transcriptome, and in
CC   detecting RNA transcripts and splice variants of human or animal
CC   transcriptomes. The libraries may also be used as specialised mini
CC   libraries to detect transcripts of a sub-transcriptome under a
CC   particular biological or pathological state, and so allowing the
CC   detection of tissue- and pathology-specific genes such as those genes
CC   only expressed in specific tissue under a specific pathological
CC   condition; to detect developmental specific genes; and to detect RNA
CC   transcripts and splice variants of a transcriptome of a patient suffering
CC   from a particular disorder. ABN27253 to ABN59589 represent
CC   oligonucleotide sequences from rats, humans and mice, which are used in
CC   the exemplification of the present invention.
CC   N.B. The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 60 BP; 16 A; 11 C; 19 G; 14 T; 0 other;

  Query Match             69.0%;  Score 13.8;  DB 24;  Length 60;
  Best Local Similarity   88.2%;  Pred. No. 2.4e+03;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 CAACATTCCAGAGCAAG 17
           ||| |||||||||||  |
Db      44 CAAAATTCCAGAGCATG 28
```

RESULT 7

```
ABN49462/c
ID   ABN49462 standard; DNA; 60 BP.
XX
AC   ABN49462;
XX
DT   15-JUL-2002  (first entry)
XX
DE   Human spliced transcript detection oligonucleotide SEQ ID NO:22210.
XX
KW   Human; mouse; rat; splice transcript; detection; RNA transcript;
KW   splice variant; transcriptome; oligonucleotide library; ss.
XX
OS   Homo sapiens.
XX
PN   WO200210449-A2.
XX
PD   07-FEB-2002.
XX
PF   20-JUL-2001; 2001WO-IB01903.
XX
PR   28-JUL-2000; 2000US-221607P.
PR   02-MAY-2001; 2001US-287724P.
XX
PA   (COMP-) COMPUGEN INC.
XX
PI   Shoshan A,  Wasserman A,  Mintz E,  Mintz L,  Faigler S;
XX
DR   WPI; 2002-257383/30.
XX
PT   New oligonucleotide libraries comprising oligonucleotides which
PT   selectively hybridize to mRNAs transcribed from a transcription unit of
PT   a genome, useful for detecting tissue-, pathology-, and
PT   developmental-specific genes  -
XX
PS   Example 1; SEQ ID 22210; 47pp; English.
XX
CC   The present invention describes oligonucleotide libraries for detecting
CC   messenger RNAs that populate a (sub-)transcriptome, where the
CC   (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC   transcription units that populate a genome. The library comprises
CC   several oligonucleotides, each capable of hybridising selectively to a
CC   set of messenger RNAs transcribed from a given transcription unit of
CC   the genome, which encodes one or more messenger RNA splice variants.
CC   The oligonucleotide libraries are useful for detecting mRNAs from a
CC   biological sample, in expression profiling studies, in qualitatively or
CC   quantitatively characterising the corresponding transcriptome, and in
CC   detecting RNA transcripts and splice variants of human or animal
CC   transcriptomes. The libraries may also be used as specialised mini
CC   libraries to detect transcripts of a sub-transcriptome under a
CC   particular biological or pathological state, and so allowing the
CC   detection of tissue- and pathology-specific genes such as those genes
CC   only expressed in specific tissue under a specific pathological
CC   condition; to detect developmental specific genes; and to detect RNA
CC   transcripts and splice variants of a transcriptome of a patient suffering
CC   from a particular disorder. ABN27253 to ABN59589 represent
CC   oligonucleotide sequences from rats, humans and mice, which are used in
CC   the exemplification of the present invention.
CC   N.B. The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 60 BP; 8 A; 15 C; 15 G; 22 T; 0 other;

  Query Match             69.0%;  Score 13.8;  DB 24;  Length 60;
  Best Local Similarity   88.2%;  Pred. No. 2.4e+03;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       3 ACATTCCAGAGCAAGGA 19
           |||||||||  | |||||
Db      51 ACATTCCAGTGAAAGGA 35
```

RESULT 8

```
ABK66079/c
ID   ABK66079 standard; DNA; 31 BP.
XX
AC   ABK66079;
XX
DT   02-JUL-2002  (first entry)
XX
DE   Human gene specific PCR primer  #167.
XX
KW   Primer; ss; DNA microarray; differential expression analysis; human.
XX
OS   Homo sapiens.
XX
PN   US6352829-B1.
XX
PD   05-MAR-2002.
XX
PF   05-JAN-1999;   99US-0225928.
XX
PR   21-MAY-1997;   97US-0859998.
XX
PA   (CLON-) CLONTECH LAB INC.
XX
PI   Chenchik A,  Jokhadze G,  Bibilashvilli R;
XX
DR   WPI; 2002-314699/35.
XX
PT   Producing sub-population of labeled nucleic acids, useful for analysing
PT   differences in RNA profiles between several different physiological
PT   sources, using set of distinct gene specific primers -
XX
PS   Example 3; SEQ ID No 167; 11pp; English.
XX
CC   The invention relates to producing a sub-population of labeled nucleic
CC   acids (NAs) comprising contacting a NA sample from a physiological
CC   source, with a pool of 50 distinct gene specific primers under suitable
CC   conditions to enzymatically generate sub-population of NAs, where
CC   each gene specific primer has a sequence complementary to a distinct
CC   mRNA, and each labeled NA is generated using a single gene specific
CC   primer. The method is useful for producing a sub-population of labeled
CC   NAs which is useful for analysing the differences in the RNA profiles
CC   between several different physiological sources, where the method
CC   comprises producing subpopulation of labeled NAs for the different
CC   physiological sources, comprising the populations for each physiological
CC   source to identify differences in the population, where the comparison
CC   is preferably performed by hybridising the labeled NAs for each of the
CC   distinct physiological sources to an array of probe NAs stably
CC   associated with the surface of a substrate to produce a hybridisation
CC   pattern for each of the sources, and comparing the patterns for each of
CC   the sources, where differential gene expression assays are
CC   utilised in differential expression analysis of diseased a normal
CC   tissue e.g. neoplastic a normal tissue, or different tissue or
CC   subtissue types. The present sequence is a human gene specific PCR
CC   primer used in the method of the invention.
CC   Note: The sequence data for this patent did not form part
CC   of the printed specification, but was obtained in electronic
CC   format directly from USPTO at
CC   http.wipo.seqdata.uspto.gov/sequence.html?DocID=6352829B1.
XX
SQ   Sequence 31 BP; 5 A; 8 C; 7 G; 11 T; 0 other;

  Query Match             68.0%;  Score 13.6;  DB 24;  Length 31;
  Best Local Similarity   80.0%;  Pred. No. 2.8e+03;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 CAACATTCCAGAGCAAGGAT 20
           || |  |||||||||| |||
Db      23 CAGCGGTCCAGAGCAATGAT 4


RESULT 9
AAQ72778/c
ID   AAQ72778 standard; DNA; 36 BP.
XX
AC   AAQ72778;
XX
DT   05-JUL-1995  (first entry)
XX
DE   Primer MPR3 for amplification of mucor rennin gene.
XX
KW   Mucor pusillus; rennin gene; primer; hybridisation; fusion gene;
KW   murine apolipoprotein E; ApoE production; treatment; hyperlipaemia;
KW   PCR; polymerase chain reaction; amplification; ss.
XX
OS   Synthetic.
XX
PN   JP06253858-A.
XX
PD   13-SEP-1994.
XX
PF   04-MAR-1993;   93JP-0069446.
XX
PR   04-MAR-1993;   93JP-0069446.
XX
PA   (BEPP/) BEPPU T.
XX
DR   WPI; 1994-329000/41.
XX
PT   Mucor rennin gene-murine apo:lipoprotein E fusion gene - for
PT   prodn. of ApoE in yeast, useful for treating hyperlipaemia
XX
PS   Example 2; Page 8; 12pp; Japanese.
XX
CC   AAQ72777-84 are primers used for the amplification of the MPR (Mucor
CC   pusillus rennin) gene, and the murine apolipoprotein E gene for use
CC   in generating a fusion gene construct. The fusion gene encodes a
CC   fusion protein which contains all or part of the prepro and mature
CC   sequence of MPR. The fused protein product is useful as an
CC   intermediate for the prepn. of murine apolipoprotein E which has the
CC   effect of lowering the cholesterol level in the plasma of a
CC   hyperlipaemia patient.
XX
SQ   Sequence 36 BP; 5 A; 8 C; 10 G; 13 T; 0 other;

  Query Match             68.0%;  Score 13.6;  DB 15;  Length 36;
  Best Local Similarity   80.0%;  Pred. No. 2.8e+03;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 CAACATTCCAGAGCAAGGAT 20
           ||||| || |||  ||||||
Db      23 CAACAGTCTAGACTAAGGAT 4


RESULT 10
AAT09079/c
ID   AAT09079 standard; DNA; 36 BP.
XX
AC   AAT09079;
XX
DT   14-MAY-1996  (first entry)
XX
DE   Mucor pusillus pepsin gene primer MPP3.
XX
KW   Microbial pepsin; Mucor pusillus; MPP; human apolipoprotein E;
KW   fusion protein; drug intermediate; PCR primer; ss.
XX
OS   Synthetic.
XX
PN   JP07241195-A.
XX
PD   19-SEP-1995.
XX
PF   04-MAR-1994;   94JP-0058269.
XX
PR   04-MAR-1994;   94JP-0058269.
XX
```

```
PA   (BEPP/) BEPPU T.
XX
DR   WPI; 1995-354276/46.
XX
PT   Fusion gene coding for Rhizomucor pepsin fused to human
PT   apo:lipoprotein E - also plasmid and transformed yeast, useful for
PT   prepn. of fusion protein as intermediate for drugs, reagents and
PT   apo:lipoprotein E
XX
PS   Example 3; Page 12; 15pp; Japanese.
XX
CC   The human apolipoprotein E gene was isolated by PCR amplification.
CC   Separately, the MPP (pepsin) gene was isolated from Mucor pusillus
CC   IFO 4578(+) and was fused to the human apolipoprotein E gene; the
CC   fusion protein expressed by the chimeric gene is useful as an
CC   intermediate for production of reagents, drugs and apolipoprotein E.
CC   The present sequence is that of a PCR primer used in the construction
CC   of an expression vector for the MPP-apoE fusion protein.
XX
SQ   Sequence 36 BP; 5 A; 8 C; 10 G; 13 T; 0 other;

  Query Match             68.0%;  Score 13.6;  DB 16;  Length 36;
  Best Local Similarity   80.0%;  Pred. No. 2.8e+03;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 CAACATTCCAGAGCAAGGAT 20
           ||||| || |||  ||||||
Db    . 23 CAACAGTCTAGACTAAGGAT 4


RESULT 11
AAZ28379/c
ID   AAZ28379 standard; DNA; 37 BP.
XX
AC   AAZ28379;
XX
DT   20-DEC-1999  (first entry)
XX
DE   PCR primer EGFPs used to amplify a green fluorescence protein EGFP.
XX
KW   PCR primer; EGFP; green fluorescence protein; quantitative analysis;
KW   gene delivery vector; retroviral packaging signal; viral RNA;
KW   gene therapy; gene modification; ss.
XX
OS   Synthetic.
XX
PN   WO9950431-A1.
XX
PD   07-OCT-1999.
XX
PF   26-MAR-1999;   99WO-AU00219.
XX
PR   27-MAR-1998;   98US-0079772.
XX
PA   (MACF-) MACFARLANE BURNET CENT MEDICAL.
XX
PI   Mak JC,  Hill MK,  Crowe SM;
XX
DR   WPI; 1999-591326/50.
XX
PT   New retroviral nucleic acid delivery vector, useful for gene
PT   modification and gene therapy -
XX
PS   Example 10; Page 37; 64pp; English.
XX
CC   PCR primers AAZ28379-Z28380 are used to amplify a 700 bp fragment of the
CC   green fluorescence protein EGFP. The PCR product is used in quantitative
CC   PCR analysis, to compare the the gene delivery efficiency of the
CC   invention with the efficiency of conventional gene delivery systems. The
CC   invention relates to a genetic construct that is useful as a nucleic
CC   acid delivery vector. The delivery vector is formed from a retroviral
CC   packaging signal, and a nucleotide sequence derived from viral RNA. The
CC   vector is in RNA form, is translatable to at least one protein and
CC   facilitates entry of a second nucleotide sequence. The invention is used
CC   in gene modification and gene therapy. The invention is more efficient
CC   than previous packaging systems.
XX
SQ   Sequence 37 BP; 11 A; 6 C; 14 G; 6 T; 0 other;

  Query Match             68.0%;  Score 13.6;  DB 20;  Length 37;
  Best Local Similarity   80.0%;  Pred. No. 2.9e+03;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 CAACATTCCAGAGCAAGGAT 20
           || ||||| |||||| | ||
Db      22 CACCATTCTAGAGCATGCAT 3


RESULT 12
AAX26807/c
ID   AAX26807 standard; DNA; 48 BP.
XX
AC   AAX26807;
XX
DT   21-JUN-1999  (first entry)
XX
DE   Primer for 583/4/5 VER/CCG mutant in HIV-1 envelope glycoproteins.
XX
KW   Site-directed mutagenesis; HIV-1 envelope glycoprotein; mutant;
KW   immune reaction; vaccine; HIV-1; PCR primer; ss.
XX
OS   Synthetic.
XX
PN   WO9916883-A2.
XX
PD   08-APR-1999.
XX
PF   01-OCT-1998;   98WO-US20693.
XX
PR   03-OCT-1997;   97US-0060808.
PR   01-OCT-1997;   97US-0060813.
XX
PA   (DAND ) DANA FARBER CANCER INST INC.
XX
PI   Farzan M,  Sodroski JG;
XX
DR   WPI; 1999-263698/22.
XX
PT   New HIV-1 proteins useful in vaccines for immunization against HIV-1
XX
PS   Disclosure; Page 34; 54pp; English.
XX
CC   PCR primers AAX26798-807 were used for site-directed mutagenesis to
CC   create novel HIV-1 envelope glycoproteins. These novel glycoproteins
CC   contain substituted cysteine residues in the alpha-helix of the gp120
CC   or gp41 transmembrane glycoprotein, and have increased coiled coil
CC   stability whilst retaining their correct folding. The novel proteins
CC   can be used to generate an immune reaction, especially as a vaccine
CC   for immunizing individuals against HIV-1.
XX
SQ   Sequence 48 BP; 12 A; 9 C; 16 G; 11 T; 0 other;

  Query Match             68.0%;  Score 13.6;  DB 20;  Length 48;
  Best Local Similarity   80.0%;  Pred. No. 2.9e+03;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 CAACATTCCAGAGCAAGGAT 20
           || | ||||| ||| |||||
Db      23 CATCCTTCCACAGCTAGGAT 4


RESULT 13
ABN36579/c
ID   ABN36579 standard; DNA; 60 BP.
XX
AC   ABN36579;
```

```
XX
DT   15-JUL-2002  (first entry)
XX
DE   Human spliced transcript detection oligonucleotide SEQ ID NO:9327.
XX
KW   Human; mouse; rat; splice transcript; detection; RNA transcript;
KW   splice variant; transcriptome; oligonucleotide library; ss.
XX
OS   Homo sapiens.
XX
PN   WO200210449-A2.
XX
PD   07-FEB-2002.
XX
PF   20-JUL-2001; 2001WO-IB01903.
XX
PR   28-JUL-2000; 2000US-221607P.
PR   02-MAY-2001; 2001US-287724P.
XX
PA   (COMP-) COMPUGEN INC.
XX
PI   Shoshan A,  Wasserman A,  Mintz E,  Mintz L,  Faigler S;
XX
DR   WPI; 2002-257383/30.
XX
PT   New oligonucleotide libraries comprising oligonucleotides which
PT   selectively hybridize to mRNAs transcribed from a transcription unit of
PT   a genome, useful for detecting tissue-, pathology-, and
PT   developmental-specific genes  -
XX
PS   Example 1; SEQ ID 9327; 47pp; English.
XX
CC   The present invention describes oligonucleotide libraries for detecting
CC   messenger RNAs that populate a (sub-)transcriptome, where the
CC   (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC   transcription units that populate a genome. The library comprises
CC   several oligonucleotides, each capable of hybridising selectively to a
CC   set of messenger RNAs transcribed from a given transcription unit of
CC   the genome, which encodes one or more messenger RNA splice variants.
CC   The oligonucleotide libraries are useful for detecting mRNAs from a
CC   biological sample, in expression profiling studies, in qualitatively or
CC   quantitatively characterising the corresponding transcriptome, and in
CC   detecting RNA transcripts and splice variants of human or animal
CC   transcriptomes. The libraries may also be used as specialised mini
CC   libraries to detect transcripts of a sub-transcriptome under a
CC   particular biological or pathological state, and so allowing the
CC   detection of tissue- and pathology-specific genes such as those genes
CC   only expressed in specific tissue under a specific pathological
CC   condition; to detect developmental specific genes; and to detect RNA
CC   transcripts and splice variants of a transcriptome of a patient suffering
CC   from a particular disorder. ABN27253 to ABN59589 represent
CC   oligonucleotide sequences from rats, humans and mice, which are used in
CC   the exemplification of the present invention.
CC   N.B. The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 60 BP; 17 A; 14 C; 12 G; 17 T; 0 other;

  Query Match             68.0%;  Score 13.6;  DB 24;  Length 60;
  Best Local Similarity   80.0%;  Pred. No. 3e+03;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 CAACATTCCAGAGCAAGGAT 20
           | | ||||||||| ||| ||
Db      51 CCATATTCCAGAGGAAGAAT 32


RESULT 14
ABA48643
ID   ABA48643 standard; DNA; 84 BP.
XX
AC   ABA48643;
XX
DT   01-FEB-2002  (first entry)
XX
DE   Human breast cell single exon nucleic acid probe #7338.
XX
KW   Human; microarray; single exon probe; gene expression; breast;
KW   disease; cancer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157271-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00662.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-496933/54.
XX
PT   New spatially-addressable set of single exon nucleic acid probes,
PT   useful for measuring gene expression in sample derived from human
PT   breast, comprises number of single exon nucleic acid probes  -
XX
PS   Claim 4; SEQ ID NO 7338; 327pp + sequence listing; English.
XX
CC   The invention relates to a spatially-addressable set of single exon
CC   nucleic acid probes for measuring gene expression in a sample derived
CC   from human breast and BT 474 cells. The method involves contacting
CC   the probes with a collection of detectably labelled nucleic acids
CC   derived from mRNA of human breast, and then measuring the label
CC   bound to each probe of the microarray. The probes are useful for
CC   verifying the expression of regions of genomic DNA predicted to
CC   encode proteins. They are useful for gene discovery, and for
CC   determining predisposition and/or prognosing breast disease. Gene
CC   expression analysis is useful for assessing the toxicity of chemical
CC   agents on cells. The microarray of this invention presents a far greater
CC   diversity of probes for measuring gene expression, with far less bias
CC   than expressed sequence tag microarrays. The method is suitable for
CC   rapid production of functional information from genomic sequence. The
CC   present sequence is a single exon nucleic acid probe of the invention.
CC   Note: The sequence data for this patent did not form part of the
CC   printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 84 BP; 33 A; 18 C; 18 G; 15 T; 0 other;

  Query Match             68.0%;  Score 13.6;  DB 22;  Length 84;
  Best Local Similarity   80.0%;  Pred. No. 3.1e+03;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 CAACATTCCAGAGCAAGGAT 20
           ||  |||||| ||||||| |
Db      47 CAGAATTCCACAGCAAGGCT 66


RESULT 15
ABA66556
ID   ABA66556 standard; DNA; 84 BP.
XX
AC   ABA66556;
XX
DT   01-FEB-2002  (first entry)
```

XX
DE Human foetal liver single exon nucleic acid probe #14861.
XX
KW Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
XX
OS Homo sapiens.
XX
PN WO200157277-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00669.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-483447/52.
XX
PT Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human fetal liver -
XX
PS Claim 4; SEQ ID NO 14861; 639pp + sequence listing; English.
XX
CC The invention relates to a single exon nucleic acid probe for
CC measuring human gene expression in a sample derived from human foetal
CC liver. The single exon nucleic acid probes may be used for predicting,
CC measuring and displaying gene expression in samples derived from human
CC fetal liver. The present sequence is a single exon nucleic acid
CC probe of the invention.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 84 BP; 33 A; 18 C; 18 G; 15 T; 0 other;

Query Match 68.0%; Score 13.6; DB 22; Length 84;
Best Local Similarity 80.0%; Pred. No. 3.1e+03;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
Qy       1 CAACATTCCAGAGCAAGGAT 20
           ||  |||||| ||||||| |
Db      47 CAGAATTCCACAGCAAGGCT 66
```

Search completed: November 23, 2002, 07:03:48
Job time : 98.1 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 26, 2002, 04:16:10 ; Search time 798.5 Seconds
(without alignments)
405.648 Million cell updates/sec

Title: US-09-296-264-24
Perfect score: 20
Sequence: 1 caacattccagagcaaggat 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 357874

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
c	1	15.4	77.0	94	14	T10880	T10880 hbc013 Huma
	2	15.2	76.0	77	9	AU260375	AU260375 AU260375
c	3	14.2	71.0	67	17	AZ514685	AZ514685 1M0361L02
	4	14.2	71.0	93	17	B33922	B33922 HS-1023-B1-
	5	13.8	69.0	72	9	AU076926	AU076926 AU076926
c	6	13.8	69.0	82	9	AI926085	AI926085 wo42b09.x
c	7	13.6	68.0	55	9	AI365089	AI365089 qt12f06.x
c	8	13.6	68.0	77	14	BQ060285	BQ060285 Ts1340 Th
	9	13.6	68.0	91	14	R82154	R82154 5E5 Chromos
	10	13.6	68.0	91	17	D86884	D86884 Human exon
	11	13.6	68.0	92	17	HSMC24E11	X88251 H.sapiens D
c	12	13.6	68.0	94	9	AA057131	AA057131 zf04f06.s
c	13	13.4	67.0	28	17	AZ312882	AZ312882 1M0029D05
c	14	13.4	67.0	63	17	AZ808041	AZ808041 2M0071M16
c	15	13.4	67.0	86	13	BM052185	BM052185 ic98b10.x
	16	13.4	67.0	100	9	AA231580	AA231580 mw34e06.r
	17	13.2	66.0	68	14	R13252	R13252 yf73a09.r1
	18	13.2	66.0	78	17	AZ491199	AZ491199 1M0324C12
c	19	13.2	66.0	85	9	AI155881	AI155881 ue07b06.r
	20	13.2	66.0	93	17	FR0018879	AL011773 F.rubripe
	21	13.2	66.0	97	9	AA063829	AA063829 mj88h10.r
c	22	12.8	64.0	25	9	AA913714	AA913714 ol38f03.s
c	23	12.8	64.0	34	9	AI035901	AI035901 ub49b03.r
	24	12.8	64.0	50	9	AU107155	AU107155 AU107155
	25	12.8	64.0	53	17	BH856261	BH856261 SALK_0810
c	26	12.8	64.0	55	17	AF107432	AF107432 AF107432
c	27	12.8	64.0	86	17	AZ775976	AZ775976 2M0009H16
	28	12.8	64.0	88	14	U25947	U25947 HSU25947 Hu
	29	12.8	64.0	96	14	T64908	T64908 yd10d11.s1
c	30	12.8	64.0	98	17	AL758257	AL758257 Arabidops
c	31	12.8	64.0	98	17	AL760317	AL760317 Arabidops
c	32	12.6	63.0	29	17	AZ596498	AZ596498 1M0409L14
	33	12.6	63.0	38	17	AL764325	AL764325 Arabidops
c	34	12.6	63.0	50	9	AU103750	AU103750 AU103750
	35	12.6	63.0	54	9	AA621269	AA621269 zu83c11.s
	36	12.6	63.0	56	9	AA634526	AA634526 zu82c11.s
	37	12.6	63.0	61	9	AI056088	AI056088 ox46f11.s
c	38	12.6	63.0	64	9	AI564913	AI564913 tm82g04.x
c	39	12.6	63.0	73	9	AA505443	AA505443 nh84b02.s
	40	12.6	63.0	73	9	AA620067	AA620067 vl53e01.r
c	41	12.6	63.0	85	9	AA286646	AA286646 vb79a01.r
	42	12.6	63.0	95	17	AZ485646	AZ485646 1M0313M08
	43	12.6	63.0	96	17	AZ762645	AZ762645 1M0557F16
	44	12.6	63.0	100	10	AW890630	AW890630 QV1-NT004
c	45	12.4	62.0	34	17	AZ810284	AZ810284 2M0074B13

ALIGNMENTS

RESULT 1
T10880/c
LOCUS T10880 94 bp mRNA linear EST 29-NOV-1993
DEFINITION hbc013 Human pancreatic islet Homo sapiens cDNA clone hbc013 3'end similar to prealbumin, mRNA sequence.
ACCESSION T10880
VERSION T10880.1 GI:391034
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 94)
AUTHORS Takeda,J., Yano,H., Eng,S., Zeng,Y. and Bell,G.I.
TITLE A molecular inventory of human pancreatic islets: sequence analysis of 1000 cDNA clones
JOURNAL Hum. Mol. Genet. 2, 1793-1798 (1993)
MEDLINE 94108427
COMMENT Contact: Bell GI or Takeda J
HHMI
Univ. of Chicago
5841 S. Maryland Ave.,MC1028,Chicago IL 60637
Tel: 3127029116
Fax: 3127020271
Email: g-bell@uchicago.edu
Seq primer: T7 primer.
FEATURES Location/Qualifiers
source 1. .94
/organism="Homo sapiens"

```
                     /db_xref="taxon:9606"
                     /clone="hbc013"
                     /clone_lib="Human pancreatic islet"
                     /note="Vector: Lambda ZAPII; Site_1: Eco RI; Site_2: Xho
                     I; mRNA was prepared from normal adult human islets. cDNA
                     was directionally synthesized from the Xho I in the vector
                     to the EcoRI site. cDNA was size fractionated to remove
                     sequences <1000 bp in size."
BASE COUNT       25 a     18 c     19 g     32 t
ORIGIN

  Query Match             77.0%;  Score 15.4;  DB 14;  Length 94;
  Best Local Similarity   94.1%;  Pred. No. 2.9e+03;
  Matches   16;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 AACATTCCAGAGCAAGG 18
           ||||||||||||  ||||
Db      18 AACATTCCAGAGAAAGG 2


RESULT 2
AU260375
LOCUS       AU260375                 77 bp    mRNA    linear   EST 25-APR-2002
DEFINITION  AU260375 3'-directed mouse cDNA library Mus musculus cDNA clone
            BED0016816 3', mRNA sequence.
ACCESSION   AU260375
VERSION     AU260375.1  GI:20327805
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 77)
  AUTHORS   Kato,K. and Matoba,R.
  TITLE     Generation of expressed sequence tags from mouse brain
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Kikuya Kato
            Graduate School of Biological Sciences
            Nara Institute of Science and Technology
            8916-5 Takayama, Ikoma, Nara 630-0101, Japan
            Tel: 81-743-72-5581
            Fax: 81-743-72-5589
            Email: kkato@bs.aist-nara.ac.jp,
            URL:http://love2.aist-nara.ac.jp/BED/index.html.
FEATURES             Location/Qualifiers
     source          1. .77
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /clone="BED0016816"
                     /clone_lib="3'-directed mouse cDNA library"
                     /tissue_type="brain"
                     /note="Vector: pGEM-T-easy"
BASE COUNT       22 a     13 c     23 g     19 t
ORIGIN

  Query Match             76.0%;  Score 15.2;  DB 9;  Length 77;
  Best Local Similarity   85.0%;  Pred. No. 3.3e+03;
  Matches   17;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 CAACATTCCAGAGCAAGGAT 20
           | ||||||||||||  |||
Db      43 CTACATTCCAGAGCAGAGAT 62


RESULT 3
AZ514685/c
LOCUS       AZ514685                 67 bp    DNA     linear   GSS 05-OCT-2000
DEFINITION  1M0361L02R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0361L02 R, DNA sequence.
ACCESSION   AZ514685
VERSION     AZ514685.1  GI:10696001
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 67)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0361  row: L  column: 02
            Seq primer: CACACAGGAAACAGCTATGACC
            Class: plasmid ends
            High quality sequence stop: 67.
FEATURES             Location/Qualifiers
     source          1. .67
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC1M0361L02"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
BASE COUNT       13 a     20 c     16 g     18 t
ORIGIN

  Query Match             71.0%;  Score 14.2;  DB 17;  Length 67;
  Best Local Similarity   84.2%;  Pred. No. 9.5e+03;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 AACATTCCAGAGCAAGGAT 20
           |||||||| | | ||||||
Db      54 AACATTCCGGTGAAAGGAT 36


RESULT 4
B33922
LOCUS       B33922                   93 bp    DNA     linear   GSS 17-OCT-1997
DEFINITION  HS-1023-B1-D09-MF.abi CIT Human Genomic Sperm Library C Homo
            sapiens genomic clone Plate=CT 802 Col=17 Row=H, DNA sequence.
ACCESSION   B33922
VERSION     B33922.1  GI:2533291
KEYWORDS    GSS.
SOURCE      human.
```

```
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 93)
  AUTHORS   Mahairas,G.G., Zackrone,K.D., Smith,T., Tipton,S., Schmidt,S.,
            Traicoff,R., Abajian,C., Blanchard,A., West,A. and Hood,L.E.
  TITLE     Construction of a Characterized Clone Resource for Genomic
            Sequencing: Generation and Preliminary Analysis of 20,000 Sequence
            Tagged Connectors
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Mahairas GG, Zackrone KD, Hood L
            University of Washington
            Seattle, WA 98195, USA
            Tel: (206) 616-8744
            Fax: (206) 685-7301
            Email: kzackron@u.washington.edu
            Sequence Tagged Connector
            Plate: CT 802  row: H  column: 17
            Class: BAC ends
            High quality sequence stop: 93.
FEATURES             Location/Qualifiers
     source          1. .93
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="Plate=CT 802 Col=17 Row=H"
                     /clone_lib="CIT Human Genomic Sperm Library C"
                     /sex="M"
                     /note="Organ: sperm; Vector: pBeloBAC11; BAC Clones in
                     E-Coli DH10B"
BASE COUNT       28 a     24 c     23 g     18 t
ORIGIN

  Query Match             71.0%;  Score 14.2;  DB 17;  Length 93;
  Best Local Similarity   84.2%;  Pred. No. 1.1e+04;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 CAACATTCCAGAGCAAGGA 19
           |||  |||||||| |||||
Db      39 CAAGCTTCCAGAGGAAGGA 57


RESULT 5
AU076926
LOCUS       AU076926                72 bp    mRNA    linear   EST 04-MAY-2000
DEFINITION  AU076926 Sugano cDNA library Homo sapiens cDNA clone kaia3454
            similar to 5'-end region of Rattus norvegicus fructose transporter
            (GLUT5) mRNA, mRNA sequence.
ACCESSION   AU076926
VERSION     AU076926.1  GI:7439458
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 72)
  AUTHORS   Suzuki,Y., Ishihara,D., Sasaki,M., Nakagawa,H., Hata,H., Tsunoda,T.
            , Watanabe,M., Komatsu,T., Ota,T., Isogai,T., Suyama,A. and Sugano
            ,S.
  TITLE     Statistical analysis of the 5' untranslated region of human mRNA
            using 'Oligo-Capped' cDNA libraries
  JOURNAL   Genomics 64 (3), 286-297 (2000)
  MEDLINE   20221373
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp
            Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano
            ,S. Construction and characterization of a full length-enriched and
            a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997)
             This clone was obtained from a 'full length-enriched' cDNA library
            constructed by 'Oligo-Capping' method. The coding region starts
            from the 50 bp upstream to the 3'-end.
FEATURES             Location/Qualifiers
     source          1. .72
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="kaia3454"
                     /clone_lib="Sugano cDNA library"
                     /note="The cDNA was prepared using the anchor primer,
                     H-T11G, from Genhunter"
BASE COUNT       20 a     17 c     23 g     12 t
ORIGIN

  Query Match             69.0%;  Score 13.8;  DB 9;  Length 72;
  Best Local Similarity   88.2%;  Pred. No. 1.5e+04;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       4 CATTCCAGAGCAAGGAT 20
           | |||||||||||| ||
Db       8 CCTTCCAGAGCAAGCAT 24


RESULT 6
AI926085/c
LOCUS       AI926085                82 bp    mRNA    linear   EST 02-SEP-1999
DEFINITION  wo42b09.x1 NCI_CGAP_Gas4 Homo sapiens cDNA clone IMAGE:2457977 3'
            similar to TR:Q92575 Q92575 MYELOBLAST KIAA0242 ;, mRNA sequence.
ACCESSION   AI926085
VERSION     AI926085.1  GI:5662049
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 82)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Seq primer: -40UP from Gibco
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .82
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2457977"
                     /clone_lib="NCI_CGAP_Gas4"
                     /tissue_type="poorly differentiated adenocarcinoma with
                     signet ring cell features"
                     /lab_host="DH10B"
                     /note="Organ: stomach; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.69 kb. Life Technologies catalog #:
                     11549-011"
BASE COUNT       21 a     16 c     19 g     26 t
ORIGIN

  Query Match             69.0%;  Score 13.8;  DB 9;  Length 82;
  Best Local Similarity   88.2%;  Pred. No. 1.6e+04;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       2 AACATTCCAGAGCAAGG 18
           ||||| ||||||||| |
Db      68 AACATCCCAGAGCAATG 52
```

```
RESULT 7
AI365089/c
LOCUS       AI365089                 55 bp    mRNA    linear   EST 13-FEB-1999
DEFINITION  qt12f06.x1 NCI_CGAP_GC4 Homo sapiens cDNA clone IMAGE:1947395 3'
            similar to SW:UBCG_HUMAN Q99462 UBIQUITIN-CONJUGATING ENZYME E2-19
            KD ;, mRNA sequence.
ACCESSION   AI365089
VERSION     AI365089.1  GI:4124778
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 55)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 902   Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .55
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1947395"
                     /clone_lib="NCI_CGAP_GC4"
                     /tissue_type="pooled germ cell tumors"
                     /lab_host="DH10B"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; 1st strand cDNA was prepared from 3 pooled
                     germ cell tumors, and was then primed with a Not I -
                     oligo(dT) primer. Double-stranded cDNA was ligated to Eco
                     RI adaptors (Pharmacia), digested with Not I and cloned
                     into the Not I and Eco RI sites of the modified pT7T3
                     vector.  Library is normalized. Library was constructed by
                     Bento Soares and M. Fatima Bonaldo. "
BASE COUNT       14 a     11 c     14 g     16 t
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 9;  Length 55;
  Best Local Similarity   80.0%;  Pred. No. 1.7e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 CAACATTCCAGAGCAAGGAT 20
           ||| | ||||| ||||||  |
Db      50 CAATAATCCAGTGCAAGGCT 31


RESULT 8
BQ060285/c
LOCUS       BQ060285                 77 bp    mRNA    linear   EST 01-APR-2002
DEFINITION  Ts1340 Thellungiella salsuginea ZAP cDNA library Thellungiella
            salsuginea cDNA similar to putative copper amine oxidase, mRNA
            sequence.
ACCESSION   BQ060285
VERSION     BQ060285.1  GI:19855234
KEYWORDS    EST.
SOURCE      Thellungiella salsuginea.
  ORGANISM  Thellungiella salsuginea
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Thellungiella.
REFERENCE   1  (bases 1 to 77)
  AUTHORS   Wang,Z.L., Li,P.H., Sun,Y.F., Zhang,Q., Zhao,Y.X. and Zhang,H.
  TITLE     Expressed sequence tags from a halophyte Thellungiella salsuginea
            cDNA library
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Hui Zhang
            Key Laboratory of Plant Stress Research
            The Biology Department of Shandong Normal University
            No.88, Wenhua East Road, Jinan, Shandong Province, 250014, PRC
            Tel: (86)531-2960864
            Fax: (86)531-2966954
            Email: zhangh@sdnu.edu.cn.
FEATURES             Location/Qualifiers
     source          1. .77
                     /organism="Thellungiella salsuginea"
                     /db_xref="taxon:72664"
                     /clone_lib="Thellungiella salsuginea ZAP cDNA library"
                     /dev_stage="seedling"
                     /note="Organ: aerial part tissue; Vector: lambda zap;
                     Site_1: EcoRI; Site_2: XhoI; total RNA extraction from
                     NaCl(200mM) treated Thellungiella salsuginea by RNAgent
                     kit(Promega); mRNA isolation by MESSAGEMAKER kit(GIBCO BRL
                     ); directional cDNA synthesis(EcoRI XhoI) by cDNA
                     synthesis kit(STRATAGEN); the ZAP express library by
                     GigapackIII Gold Cloning kit(STRATAGENE)"
BASE COUNT       20 a     14 c     21 g     22 t
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 14;  Length 77;
  Best Local Similarity   80.0%;  Pred. No. 1.9e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 CAACATTCCAGAGCAAGGAT 20
           ||||| | || || ||||||
Db      60 CAACAGTACATAGAAAGGAT 41


RESULT 9
R82154
LOCUS       R82154                   91 bp    mRNA    linear   EST 21-SEP-1995
DEFINITION  5E5 Chromosome 21 exon Homo sapiens cDNA 5' and 3', mRNA sequence.
ACCESSION   R82154
VERSION     R82154.1  GI:994958
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 91)
  AUTHORS   McCormick,M.K.
  TITLE     Chromosome 21 exons
  JOURNAL   Unpublished (1995)
COMMENT     Contact: McCormick MK
            Molecular Neurogenetics Unit
            Massachusetts General Hospital
            MGH East, Building 149, 13th St., Charlestown MA 02129
            Tel: 6177249616
            Fax: 6177265736
            Email: McCormick@helix.MGH.Harvard.edu
            Seq primer: T3 and T7.
FEATURES             Location/Qualifiers
     source          1. .91
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="Chromosome 21 exon"
                     /lab_host="E. coli DH5a"
                     /note="Vector: pBluescriptIIKS+; Exons were isolated from
                     human chromosome 21 specific cosmids  using a modification
                     of the method of exon amplification (Proc. Natl. Acad.
                     Sci. USA 88:4005-4009, 1991).   Amplified exons were
                     digested with SalI and BamHI and subsequently cloned into
                     pBluescriptIIKS+ at the SalI and BglII sites."
```

```
BASE COUNT       26 a     30 c     21 g     14 t
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 14;  Length 91;
  Best Local Similarity   80.0%;  Pred. No. 2.1e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 CAACATTCCAGAGCAAGGAT 20
          |||| | |||| || |||||
Db     47 CAACCTGCCAGTGCCAGGAT 66


RESULT 10
D86884
LOCUS       D86884                   91 bp    DNA     linear   GSS 07-FEB-1999
DEFINITION  Human exon sequence in 1.6Mb segment encompassing Down's syndrome
            region, exon, genomic survey sequence.
ACCESSION   D86884
VERSION     D86884.1  GI:1813398
KEYWORDS    GSS; exon trapping for Down's syndrome region.
SOURCE      Homo sapiens DNA, clone:E19-12.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (sites)
  AUTHORS   Ohira,M., Seki,N., Nagase,T., Suzuki,E., Nomura,N., Ohara,O.,
            Hattori,M., Sakaki,Y., Eki,T., Murakami,Y., Saito,T., Ichikawa,H.
            and Ohki,M.
  TITLE     Gene identification in the 1.6 -Mb of the down syndrome region on
            chromosome 21
  JOURNAL   Genome Res. (1997) In press
REFERENCE   2  (sites)
  AUTHORS   Ohira,M., Seki,N., Nagase,T., Ichikawa,H., Suzuki,E., Nomura,N. and
            Ohki,M.
  TITLE     Gene Identification in a 1.6-Mb Region of the Down Syndrome Region
            on Chromosome 21
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 91)
  AUTHORS   Ohira,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-AUG-1996) Miki Ohira, Kazusa DNA Research Institute,
            Laboratory of Gene Structure 1; 1532-3 Yanauchino, Kisarazu, Chiba
            292, Japan (E-mail:oohira@kazusa.or.jp, Tel:+81-438-52-3932,
            Fax:+81-438-52-3931)
FEATURES             Location/Qualifiers
     source          1. .91
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="21"
                     /map="21q22.2"
                     /clone="E19-12"
     exon            1. .91
                     /note="trapped exon sequence"
BASE COUNT       26 a     30 c     21 g     14 t
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 17;  Length 91;
  Best Local Similarity   80.0%;  Pred. No. 2.1e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 CAACATTCCAGAGCAAGGAT 20
          |||| | |||| || |||||
Db     47 CAACCTGCCAGTGCCAGGAT 66


RESULT 11
HSMC24E11
LOCUS       HSMC24E11                92 bp    DNA     linear   GSS 29-MAY-1997
DEFINITION  H.sapiens DNA for trapped exon (ID HMC24E11), genomic survey
            sequence.
ACCESSION   X88251
VERSION     X88251.1  GI:1437808
KEYWORDS    GSS.
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 92)
  AUTHORS   Chen,H.M., Rossier,C., Chrast,R. and Antonarakis,S.E.
  TITLE     Cloning of trapped exons from human chromosome 21
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 92)
  AUTHORS   Antonarakis,S.E.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-MAR-1995) Stylianos E. Antonarakis, Division of
            Medical Genetics, University and Cantonal Hospital of Geneva, CMU,
            1 rue Michel-Servet, 1211 Geneva, SWITZERLAND
REFERENCE   3  (bases 1 to 92)
  AUTHORS   Chen,H., Chrast,R., Rossier,C., Morris,M.A., Lalioti,M.D. and
            Antonarakis,S.E.
  TITLE     Cloning of 559 potential exons of genes of human chromosome 21 by
            exon trapping
  JOURNAL   Genome Res. 6 (8), 747-760 (1996)
  MEDLINE   97011340
   PUBMED   8858350
FEATURES             Location/Qualifiers
     source          1. .92
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="21"
     exon            1. .92
                     /note="trapped exon"
BASE COUNT       26 a     30 c     21 g     14 t      1 others
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 17;  Length 92;
  Best Local Similarity   80.0%;  Pred. No. 2.1e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 CAACATTCCAGAGCAAGGAT 20
          |||| | |||| || |||||
Db     48 CAACCTGCCAGTGCCAGGAT 67


RESULT 12
AA057131/c
LOCUS       AA057131                 94 bp    mRNA    linear   EST 01-FEB-1997
DEFINITION  zf04f06.s1 Soares_fetal_heart_NbHH19W Homo sapiens cDNA clone
            IMAGE:375971 3' similar to SW:LEG7_HUMAN P47929 GALECTIN-7 ;, mRNA
            sequence.
ACCESSION   AA057131
VERSION     AA057131.1  GI:1549652
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 94)
  AUTHORS   Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiapelli,B.,
            Chissoe,S., Dietrich,N., DuBuque,T., Favello,A., Gish,W., Hawkins
            ,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N., Mardis,E., Moore
            ,B., Morris,M., Parsons,J., Prange,C., Rifkin,L., Rohlfing,T.,
            Schellenberg,K., Soares,M.B., Tan,F., Thierry-Meg,J., Trevaskis,E.,
            Underwood,K., Wohldmann,P., Waterston,R., Wilson,R. and Marra,M.
  TITLE     Generation and analysis of 280,000 human expressed sequence tags
  JOURNAL   Genome Res. 6 (9), 807-828 (1996)
  MEDLINE   97044478
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Trace considered overall poor quality
```

```
            Possible reversed clone: similarity on wrong strand
            Insert Length: 425   Std Error: 0.00
            Seq primer: -40M13 fwd. from Amersham
            High quality sequence stop: 1.
FEATURES              Location/Qualifiers
     source           1. .94
                      /organism="Homo sapiens"
                      /db_xref="GDB:1284227"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:375971"
                      /clone_lib="Soares_fetal_heart_NbHH19W"
                      /sex="unknown"
                      /dev_stage="19 weeks"
                      /lab_host="DH10B (ampicillin resistant)"
                      /note="Organ: heart; Vector: pT7T3D (Pharmacia) with a
                      modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                      strand cDNA was primed with a Not I - oligo(dT) primer [5'
                      TGTTACCAATCTGAAGTGGGAGCGGCCGCATCTTTTTTTTTTTTTTTTT 3'],
                      double-stranded cDNA was size selected, ligated to Eco RI
                      adapters (Pharmacia), digested with Not I and cloned into
                      the Not I and Eco RI sites of a modified pT7T3 vector
                      (Pharmacia). Library went through one round of
                      normalization to a Cot = 5. Library constructed by
                      M.Fatima Bonaldo. This library was constructed from the
                      same fetus as the fetal lung library, Soares fetal lung
                      NbHL19W."
BASE COUNT       14 a     30 c     26 g     21 t      3 others
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 9;  Length 94;
  Best Local Similarity   80.0%;  Pred. No. 2.1e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 CAACATTCCAGAGCAAGGAT 20
          |||||   |||||||||| |
Db     26 CAACAGCACAGAGCAAGGCT 7


RESULT 13
AZ312882/c
LOCUS       AZ312882                 28 bp    DNA     linear   GSS 29-SEP-2000
DEFINITION  1M0029D05F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0029D05 F, DNA sequence.
ACCESSION   AZ312882
VERSION     AZ312882.1  GI:10357254
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 28)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0029  row: D  column: 05
            Seq primer: CGTTGTAAAACGACGGCCAGT
            Class: plasmid ends
            High quality sequence stop: 28.
FEATURES              Location/Qualifiers
     source           1. .28
                      /organism="Mus musculus"
                      /strain="C57BL/6J"
                      /db_xref="taxon:10090"
                      /clone="UUGC1M0029D05"
                      /clone_lib="Mouse 10kb plasmid UUGC1M library"
                      /sex="Male"
                      /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                      /note="Vector: PWD42nv; Purified genomic DNA from M.
                      musculus C57BL/6J (male) was obtained  from the Jackson
                      Laboratory Mouse DNA Resource
                      (http://www.jax.org/resources/documents/dnares/). The DNA
                      was hydrodynamically sheared by repeated passage through a
                      0.005 inch orifice at constant velocity. The sheared DNA
                      was blunt end-repaired with T4 DNA polymerase and T4
                      polynucleotide kinase. Adaptor oligonucleotides were
                      ligated to the blunt ends in high  molar excess. The
                      adaptored DNA was purified and size-selected for a 9.5 to
                      10.5 kb range using preparative agarose gel
                      electrophoresis. Vector DNA was prepared from a derivative
                      of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                      inducible derivative of plasmid R1. The vector was ligated
                      with adaptors complementary to the insert adaptors and
                      purified. The sheared, adaptored mouse DNA was annealed to
                      adaptored vector DNA, and transformed into
                      chemically-competent E. coli XL10-Gold (Stratagene) cells
                      and selected for ampicillin resistance."
BASE COUNT       6 a     10 c      3 g      9 t
ORIGIN

  Query Match             67.0%;  Score 13.4;  DB 17;  Length 28;
  Best Local Similarity   93.3%;  Pred. No. 1.6e+04;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      5 ATTCCAGAGCAAGGA 19
          ||||||||| |||||
Db     24 ATTCCAGAGTAAGGA 10


RESULT 14
AZ808041/c
LOCUS       AZ808041                 63 bp    DNA     linear   GSS 20-FEB-2001
DEFINITION  2M0071M16F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC2M0071M16 F, DNA sequence.
ACCESSION   AZ808041
VERSION     AZ808041.1  GI:12972989
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 63)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0071  row: M  column: 16
            Seq primer: CGTTGTAAAACGACGGCCAGT
            Class: plasmid ends
            High quality sequence stop: 63.
FEATURES              Location/Qualifiers
     source           1. .63
```

/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC2M0071M16"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptored DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptored mouse DNA was annealed to adaptored vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."
BASE COUNT 11 a 14 c 18 g 20 t
ORIGIN

Query Match 67.0%; Score 13.4; DB 17; Length 63;
Best Local Similarity 93.3%; Pred. No. 2.2e+04;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 4 CATTCCAGAGCAAGG 18
||| |||||||||||
Db 49 CATACCAGAGCAAGG 35

RESULT 15
BM052185/c
LOCUS BM052185 86 bp mRNA linear EST 12-MAR-2002
DEFINITION ic98b10.x3 Melton Normalized Mixed Mouse Pancreas 1 N1-MMS1 Mus musculus cDNA clone IMAGE:5661643 3', mRNA sequence.
ACCESSION BM052185
VERSION BM052185.1 GI:16807372
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 86)
AUTHORS Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K., Lemishka,I., Sceare,M., Brestelli,J., Gradwohl,G., Clifton,S., Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A., Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J., Cardenas ,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R., Williams,T. , Jackson,Y. and Bowers,Y.
TITLE Endocrine Pancreas Consortium
JOURNAL Unpublished (2000)
COMMENT Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge, MA 02138
Tel: 617-495-1812
Fax: 617-495-8557
Email: dmelton@biohp.harvard.edu
Library was constructed by Dr. Douglas Melton DNA sequencing by: Washington University Genome Sequencing Center For information on obtaining a clone please contact: Juliana Brown (brown@fas.harvard.edu)
MGI:1947969 This sequence now available from the IMAGE consortium, for clone orders contact: info@image.llnl.gov.
FEATURES Location/Qualifiers
source 1. .86
/organism="Mus musculus"
/strain="ICR"
/db_xref="taxon:10090"
/clone="IMAGE:5661643"
/clone_lib="Melton Normalized Mixed Mouse Pancreas 1 N1-MMS1"
/sex="Both for embryonic & newborn, male for adult and adult islet"
/dev_stage="Embryonic day 10.5, E12.5, E16.5, newborn, adult, mixed"
/lab_host="DH10B"
/note="Vector: pSPORT1; Site_1: Not I; Site_2: Sal I; Five libraries representing EI0.5/12.5 pancreatic bud, E16.5 pancreas, newborn pancreas, adult pancreas, and adult islets of Langerhans were seperately constructed using SuperScript Plasmid Library kit (Life Technologies). cDNA was made by oligo-dT priming and size-selected by column fractionation. Libraries were amplified once on solid support and plasmid DNA from each library was prepared and mixed in equal amounts. The mixed library DNA was normalized by method #4 from Bonaldo, Lennon, and Soares 1996 Genome Research 6:791-806; 0.5 microgram single-stranded mixed library plasmid DNA was mixed with 5 micrograms PCR product representing mixed library inserts and hybridized to an Ecot of 6. Single-stranded (unhybridized) plasmids were isolated by hydroxyapatite chromatography and used to make this library."
BASE COUNT 19 a 9 c 29 g 29 t
ORIGIN

Query Match 67.0%; Score 13.4; DB 13; Length 86;
Best Local Similarity 93.3%; Pred. No. 2.5e+04;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 AACATTCCAGAGCAA 16
||||||| |||||||
Db 61 AACATTCAAGAGCAA 47

Search completed: November 26, 2002, 17:57:41
Job time : 808.5 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 3, 2002, 18:14:22 ; Search time 351.3 Seconds
(without alignments)
1656.863 Million cell updates/sec

Title: US-09-296-264-25
Perfect score: 20
Sequence: 1 cgatcttgaacttcctcatg 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 995600

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
c 1	14.4	72.0	94	11	G71345	G71345 682726431FM
2	14.2	71.0	34	6	AX101195	AX101195 Sequence
3	13.8	69.0	18	6	I69019	I69019 Sequence 28
c 4	13.8	69.0	27	6	I69018	I69018 Sequence 28
5	13.8	69.0	60	6	AR031477	AR031477 Sequence
c 6	13.8	69.0	60	6	AR031477	AR031477 Sequence
c 7	13.8	69.0	76	8	AF317969	AF317969 Arabidops
c 8	13.8	69.0	85	10	MMHIF3A04	AF079143 Mus muscu
9	13.6	68.0	36	6	AX377543	AX377543 Sequence
c 10	13.6	68.0	46	6	AX382066	AX382066 Sequence
c 11	13.4	67.0	27	6	AX101025	AX101025 Sequence
12	13.4	67.0	35	6	E35118	E35118 Truncated c
13	13.2	66.0	24	6	AX092245	AX092245 Sequence
14	13.2	66.0	25	6	A78571	A78571 Sequence 3
15	13.2	66.0	25	6	I68074	I68074 Sequence 3
c 16	13.2	66.0	26	6	AR130849	AR130849 Sequence
c 17	13.2	66.0	42	6	A30247	A30247 Human beta-
c 18	13.2	66.0	42	6	AR001607	AR001607 Sequence
c 19	13.2	66.0	50	6	AX199432	AX199432 Sequence
c 20	13.2	66.0	51	6	AX199431	AX199431 Sequence
21	13.2	66.0	54	6	A34826	A34826 pSHVG21 lin
c 22	13.2	66.0	57	7	ECBCGENE	X82851 Escherichia
23	13.2	66.0	62	6	A14802	A14802 Nucleotide
c 24	13.2	66.0	62	6	A14803	A14803 Nucleotide
c 25	13.2	66.0	65	6	AX485443	AX485443 Sequence
c 26	13.2	66.0	70	1	ECHUPB3	X54554 E.coli hupB
27	13.2	66.0	80	6	E31016	E31016 Method for
28	13.2	66.0	88	4	BTU89646	U89646 Bos taurus
29	12.8	64.0	19	6	AX130977	AX130977 Sequence
30	12.8	64.0	19	6	AX130978	AX130978 Sequence
31	12.8	64.0	19	6	AX130979	AX130979 Sequence
32	12.8	64.0	25	6	AR148536	AR148536 Sequence
33	12.8	64.0	25	6	AR148542	AR148542 Sequence
34	12.8	64.0	25	6	I62315	I62315 Sequence 18
35	12.8	64.0	25	6	I62321	I62321 Sequence 24
36	12.8	64.0	25	6	I62331	I62331 Sequence 34
c 37	12.8	64.0	31	6	AX248637	AX248637 Sequence
38	12.8	64.0	47	6	I24533	I24533 Sequence 13
39	12.8	64.0	47	6	I33875	I33875 Sequence 13
c 40	12.8	64.0	55	9	HSDSRAD1C	X98561 H.sapiens m
41	12.8	64.0	72	6	I76362	I76362 Sequence 32
c 42	12.8	64.0	73	6	AR012412	AR012412 Sequence
c 43	12.8	64.0	73	6	AR020240	AR020240 Sequence
c 44	12.8	64.0	73	6	AR109261	AR109261 Sequence
c 45	12.8	64.0	73	6	I82586	I82586 Sequence 27

ALIGNMENTS

RESULT 1
G71345/c
LOCUS G71345 94 bp DNA linear STS 08-JUN-2001
DEFINITION 682726431FMo17 maize leaf DNA Zea mays STS genomic, sequence tagged site.
ACCESSION G71345
VERSION G71345.1 GI:14333030
KEYWORDS STS.
SOURCE Zea mays.
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 94)
AUTHORS Yang,Y.J., Guo,L., Ashlock,D.A., Wen,T.J. and Schnable,P.S.

```
  TITLE     3' UTR sequences of maize genes
  JOURNAL   Unpublished (2001)
COMMENT
            Contact: Schnable, P.S.
            Schnable laboratory
            Iowa State University
            G405 Agronomy Hall, Ames, IA 50011, USA
            Tel: 515-294-0975
            Fax: 515-294-2299
            Email: schnable@iastate.edu
            Primer A: GGCAGCTGGGTTCCATC
            Primer B: GTGCATCTTGCATTACCTGC
            PCR Profile:
                 Denaturation:   94 degrees C for 30 seconds
                 Annealing:      60 degrees C for 45 seconds
                 Polymerization: 72 degrees C for 90 seconds
                 PCR cycles:     31
                 Thermal cycler: Perkin Elmer TC
            Protocol:
                   Template: 10-20 ng
                   Primer: each 0.5 uM
                   dNTPs: each 200 uM
                   Taq Polymerase: 0.05 units/ul
                   Total vol: 20 ul

            Buffer:
                  MgCl2: 2 mM
                  KCl: 50 mM
                  Tris-HCl: 20 mM
                  pH: 8.4.
FEATURES             Location/Qualifiers
     source          1. .94
                     /organism="Zea mays"
                     /strain="DE811"
                     /db_xref="taxon:4577"
                     /clone_lib="maize leaf DNA"
                     /note="PCR products amplified from genomic DNA"
     STS             <1. .>94
BASE COUNT       31 a     22 c     31 g     10 t
ORIGIN

  Query Match             72.0%;  Score 14.4;  DB 11;  Length 94;
  Best Local Similarity   93.8%;  Pred. No. 6.4e+03;
  Matches   15;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 CGATCTTGAACTTCCT 16
           |||||||||| ||||||
Db      52 CGATCTTGAGCTTCCT 37


RESULT 2
AX101195
LOCUS       AX101195                  34 bp    DNA     linear   PAT 10-APR-2001
DEFINITION  Sequence 7 from Patent WO0121808.
ACCESSION   AX101195
VERSION     AX101195.1  GI:13620017
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 34)
  AUTHORS   Duwat,P.D., le Loir,Y. and Gaudu,P.
  TITLE     Lactic acid bacteria transformed to be provided with respiratory
            metabolism, and ferments comprising said lactic acid bacteria
  JOURNAL   Patent: WO 0121808-A 7 29-MAR-2001;
            INSTITUT NATIONAL DE LA RECHERCHE AGRONOMIQUE (INRA) (FR;)
FEATURES             Location/Qualifiers
     source          1. .34
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="amorce d'amplification"
BASE COUNT        9 a      8 c      2 g     15 t
ORIGIN

  Query Match             71.0%;  Score 14.2;  DB 6;  Length 34;
  Best Local Similarity   84.2%;  Pred. No. 6.8e+03;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 CGATCTTGAACTTCCTCAT 19
           ||||||| ||| |||| ||
Db       5 CGATCTTTAACGTCCTAAT 23


RESULT 3
I69019
LOCUS       I69019                    18 bp    DNA     linear   PAT 04-FEB-1998
DEFINITION  Sequence 289 from patent US 5677149.
ACCESSION   I69019
VERSION     I69019.1  GI:2831141
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 18)
  AUTHORS   Bauer,S.Christopher., Abrams,M.Allen., Braford-Goldberg,S.Ruth.,
            Caparon,M.Helena., Easton,A.Michael., Klein,B.Kure.,
            McKearn,J.Patrick., Olins,P., Paik,K., Polazzi,J. and
            Thomas,J.Warren.
  TITLE     Interleukin-3 (IL-3) mutant polypeptides and their recombinant
            production
  JOURNAL   Patent: US 5677149-A 289 14-OCT-1997;
FEATURES             Location/Qualifiers
     source          1. .18
                     /organism="unknown"
BASE COUNT        3 a      6 c      2 g      7 t
ORIGIN

  Query Match             69.0%;  Score 13.8;  DB 6;  Length 18;
  Best Local Similarity   88.2%;  Pred. No. 1e+04;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       3 ATCTTGAACTTCCTCAT 19
           ||||||| ||||| |||
Db       1 ATCTTGAGCTTCCCCAT 17


RESULT 4
I69018/c
LOCUS       I69018                    27 bp    DNA     linear   PAT 04-FEB-1998
DEFINITION  Sequence 288 from patent US 5677149.
ACCESSION   I69018
VERSION     I69018.1  GI:2831140
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 27)
  AUTHORS   Bauer,S.Christopher., Abrams,M.Allen., Braford-Goldberg,S.Ruth.,
            Caparon,M.Helena., Easton,A.Michael., Klein,B.Kure.,
            McKearn,J.Patrick., Olins,P., Paik,K., Polazzi,J. and
            Thomas,J.Warren.
  TITLE     Interleukin-3 (IL-3) mutant polypeptides and their recombinant
            production
  JOURNAL   Patent: US 5677149-A 288 14-OCT-1997;
FEATURES             Location/Qualifiers
     source          1. .27
                     /organism="unknown"
BASE COUNT       11 a      6 c      6 g      4 t
ORIGIN

  Query Match             69.0%;  Score 13.8;  DB 6;  Length 27;
  Best Local Similarity   88.2%;  Pred. No. 1.1e+04;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       3 ATCTTGAACTTCCTCAT 19
           ||||||| ||||| |||
```

```
Db     27 ATCTTGAGCTTCCCCAT 11


RESULT 5
AR031477
LOCUS       AR031477                 60 bp    DNA     linear   PAT 29-SEP-1999
DEFINITION  Sequence 34 from patent US 5866363.
ACCESSION   AR031477
VERSION     AR031477.1  GI:5945766
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 60)
  AUTHORS   Pieczenik,G.
  TITLE     Method and means for sorting and identifying biological information
  JOURNAL   Patent: US 5866363-A 34 02-FEB-1999;
FEATURES             Location/Qualifiers
     source          1. .60
                     /organism="unknown"
BASE COUNT       19 a     11 c     11 g     19 t
ORIGIN

  Query Match            69.0%;  Score 13.8;  DB 6;  Length 60;
  Best Local Similarity  88.2%;  Pred. No. 1.3e+04;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      4 TCTTGAACTTCCTCATG 20
          |||||||  ||||||||
Db     16 TCTTGAAGATCCTCATG 32


RESULT 6
AR031477/c
LOCUS       AR031477                 60 bp    DNA     linear   PAT 29-SEP-1999
DEFINITION  Sequence 34 from patent US 5866363.
ACCESSION   AR031477
VERSION     AR031477.1  GI:5945766
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 60)
  AUTHORS   Pieczenik,G.
  TITLE     Method and means for sorting and identifying biological information
  JOURNAL   Patent: US 5866363-A 34 02-FEB-1999;
FEATURES             Location/Qualifiers
     source          1. .60
                     /organism="unknown"
BASE COUNT       19 a     11 c     11 g     19 t
ORIGIN

  Query Match            69.0%;  Score 13.8;  DB 6;  Length 60;
  Best Local Similarity  88.2%;  Pred. No. 1.3e+04;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      4 TCTTGAACTTCCTCATG 20
          |||||||  ||||||||
Db     45 TCTTGAAGATCCTCATG 29


RESULT 7
AF317969/c
LOCUS       AF317969                 76 bp    DNA     linear   PLN 01-MAR-2001
DEFINITION  Arabidopsis thaliana small nucleolar RNA R45, complete sequence.
ACCESSION   AF317969
VERSION     AF317969.1  GI:13172755
KEYWORDS    .
SOURCE      Arabidopsis thaliana.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 76)
  AUTHORS   Barneche,F., Gaspin,C., Guyot,R. and Echeverria,M.
  TITLE     Extensive characterization of methylation guide small nucleolar
            RNAs in Arabidopsis thaliana
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 76)
  AUTHORS   Barneche,F., Gaspin,C., Guyot,R. and Echeverria,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-NOV-2000) Laboratoire Genome et Developement des
            Plantes, UMR CNRS 5096, Universite de Perpignan, Avenue de
            Villeneuve, Perpignan 66100, France
FEATURES             Location/Qualifiers
     source          1. .76
                     /organism="Arabidopsis thaliana"
                     /db_xref="taxon:3702"
                     /dev_stage="seedling"
                     /note="ecotype: Columbia O"
     misc_RNA        1. .76
                     /product="small nucleolar RNA R45"
                     /note="AtsnoR45"
                     /evidence=experimental
BASE COUNT       25 a      9 c     18 g     24 t
ORIGIN

  Query Match            69.0%;  Score 13.8;  DB 8;  Length 76;
  Best Local Similarity  88.2%;  Pred. No. 1.3e+04;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      3 ATCTTGAACTTCCTCAT 19
          |||||||||  || ||||
Db     22 ATCTTGAACATCATCAT 6


RESULT 8
MMHIF3A04/c
LOCUS       MMHIF3A04                85 bp    DNA     linear   ROD 06-NOV-2001
DEFINITION  Mus musculus hypoxia-inducible factor 3 alpha gene, exon 4.
ACCESSION   AF079143
VERSION     AF079143.1  GI:6650260
KEYWORDS    .
SEGMENT     4 of 14
SOURCE      Mus musculus.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 85)
  AUTHORS   Gu,Y.Z., Moran,S.M., Hogenesch,J.B., Wartman,L. and Bradfield,C.A.
  TITLE     Molecular characterization and chromosomal localization of a third
            alpha-class hypoxia inducible factor subunit, HIF3alpha
  JOURNAL   Gene Expr. 7 (3), 205-213 (1998)
  MEDLINE   99054547
   PUBMED   9840812
REFERENCE   2  (bases 1 to 85)
  AUTHORS   Gu,Y.-Z., Moran,S.M., Hogenesch,J.B., Wartman,L. and Bradfield,C.A.
  TITLE     Cloning and Characterization of a Third Hypoxia Inducible Factor,
            HIF3-alpha
  JOURNAL   J. Biol. Chem. (1999) In press
REFERENCE   3  (bases 1 to 85)
  AUTHORS   Gu,Y.-Z., Moran,S.M., Hogenesch,J.B., Wartman,L. and Bradfield,C.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-JUL-1998) McArdle Cancer Center, University of
            Wisconsin, 1400 University Avenue, Madison, WI 53706, USA
FEATURES             Location/Qualifiers
     source          1. .85
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
     exon            1. .85
                     /number=4
BASE COUNT       21 a     26 c     18 g     20 t
ORIGIN

  Query Match            69.0%;  Score 13.8;  DB 10;  Length 85;
  Best Local Similarity  88.2%;  Pred. No. 1.3e+04;
```

```
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      4 TCTTGAACTTCCTCATG 20
          ||||||| |||||| ||
Db     65 TCTTGAAGTTCCTCTTG 49


RESULT 9
AX377543
LOCUS       AX377543                 36 bp    DNA     linear   PAT 18-MAR-2002
DEFINITION  Sequence 20 from Patent WO0212553.
ACCESSION   AX377543
VERSION     AX377543.1  GI:19573729
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Kappel,A., Polakowski,T., Pignot,M., Windhab,N., Behrensdorf,H. and
            Muth,J.
  TITLE     Method for detecting mutations in nucleotide sequences
  JOURNAL   Patent: WO 0212553-A 20 14-FEB-2002;
            Nanogen Recognomics GmbH (DE)
FEATURES             Location/Qualifiers
     source          1. .36
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="Oligonucleotid 'TT'"
BASE COUNT        7 a      9 c      8 g     12 t
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 6;  Length 36;
  Best Local Similarity   80.0%;  Pred. No. 1.5e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 CGATCTTGAACTTCCTCATG 20
          || |||| |||||||  |||
Db     16 CGCTCTTTAACTTCCGTATG 35


RESULT 10
AX382066/c
LOCUS       AX382066                 46 bp    DNA     linear   PAT 18-MAR-2002
DEFINITION  Sequence 10 from Patent WO0166773.
ACCESSION   AX382066
VERSION     AX382066.1  GI:19576884
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Bryan,G.T., Mcgonigle,B., Maxwell,C.A., Potter,S.M. and Hwang,D.C.
  TITLE     Nucleic acids that code for oxidosqualene cyclases
  JOURNAL   Patent: WO 0166773-A 10 13-SEP-2001;
            E.I. DUPONT DE NEMOURS AND COMPANY (US)
FEATURES             Location/Qualifiers
     source          1. .46
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="PCR primer"
BASE COUNT       18 a      8 c     13 g      7 t
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 6;  Length 46;
  Best Local Similarity   80.0%;  Pred. No. 1.5e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 CGATCTTGAACTTCCTCATG 20
          | ||||| ||| ||| ||||
Db     43 CTATCTTTAACCTCCACATG 24


RESULT 11
AX101025/c
LOCUS       AX101025                 27 bp    DNA     linear   PAT 10-APR-2001
DEFINITION  Sequence 5 from Patent WO0121653.
ACCESSION   AX101025
VERSION     AX101025.1  GI:13619881
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 27)
  AUTHORS   Dillon,D.C.
  TITLE     Ovarian tumor antigen and methods of use therefor
  JOURNAL   Patent: WO 0121653-A 5 29-MAR-2001;
            CORIXA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .27
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="HPP14 primer sequence"
BASE COUNT       12 a      5 c      5 g      5 t
ORIGIN

  Query Match             67.0%;  Score 13.4;  DB 6;  Length 27;
  Best Local Similarity   93.3%;  Pred. No. 1.8e+04;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      2 GATCTTGAACTTCCT 16
          ||||||||||||| |
Db     26 GATCTTGAACTTCTT 12


RESULT 12
E35118
LOCUS       E35118                   35 bp    DNA     linear   PAT 18-JUN-2001
DEFINITION  Truncated cellulase composition.
ACCESSION   E35118
VERSION     E35118.1  GI:13018943
KEYWORDS    JP 1999221086-A/20.
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 35)
  AUTHORS   Paiji,A., Petaer,L.B., Roy,M.D., Gurahamu,K.F., Moreland,D.G.,
            Hyu,M. and Daian,P.W.
  TITLE     Truncated cellulase composition
  JOURNAL   Patent: JP 1999221086-A 20 17-AUG-1999;
            CLARIANT INTERNATIONAL LTD
COMMENT     OS   Artificial Sequence
            PN   JP 1999221086-A/20
            PD   17-AUG-1999
            PF   21-SEP-1998 JP 1998283606
            PR   19-SEP-1997 US    08/932571
            PI   PAIJI ANDERSON,PETAER  L  BAGUKUISUTO,ROY  M DANIEL, PI
            GURAHAMU K FARINTON,
            PI   MORELAND DAVID GIBUSU,HYU MORGAN,DAIAN PURATONIOTISU WILLIAM
            PC   C12N15/09,C11D3/386,C12N1/21,C12N9/42//(C12N1/21,C12R1:19), PC
              (C12N9/42,C12R1:19),C12N15/00
            CC
            FH   Key             Location/Qualifiers
            FT   source          1. .35
            FT                   /organism='Artificial Sequence'.
FEATURES             Location/Qualifiers
     source          1. .35
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
BASE COUNT       10 a      7 c      6 g     12 t
ORIGIN

  Query Match             67.0%;  Score 13.4;  DB 6;  Length 35;
  Best Local Similarity   93.3%;  Pred. No. 1.9e+04;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      6 TTGAACTTCCTCATG 20
```

```
          |||||||||| ||||
Db      1 TTGAACTTCCCCATG 15


RESULT 13
AX092245
LOCUS       AX092245                  24 bp    DNA     linear   PAT 21-MAR-2001
DEFINITION  Sequence 24 from Patent WO0116171.
ACCESSION   AX092245
VERSION     AX092245.1  GI:13444444
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 24)
  AUTHORS   Mallet,F., Cosset,F.L., Blond,J.L., Lavillette,D., Bouton,O. and
            Ruggieri,A.
  TITLE     Method for detecting the expression of an envelope protein of a
            human endogenous retrovirus and uses of a gene coding for said
            protein
  JOURNAL   Patent: WO 0116171-A 24 08-MAR-2001;
            BIO MERIEUX (FR) ; INSTITUT NATIONAL DE LA SANTE ET DE LA RECHERCHE
            MEDICALE (INSERM) (FR)
FEATURES             Location/Qualifiers
     source          1. .24
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT        6 a      5 c      3 g     10 t
ORIGIN

  Query Match            66.0%;  Score 13.2;  DB 6;  Length 24;
  Best Local Similarity  83.3%;  Pred. No. 2.2e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 CGATCTTGAACTTCCTCA 18
          |||||| ||| ||| |||
Db      5 CGATCTGGAATTTCTTCA 22


RESULT 14
A78571
LOCUS       A78571                    25 bp    DNA     linear   PAT 19-OCT-1999
DEFINITION  Sequence 3 from Patent EP0576862.
ACCESSION   A78571
VERSION     A78571.1  GI:6090232
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 25)
  AUTHORS   Abken,H.J. and Barchet,H.
  TITLE     DNA FOR IMMORTALISATION OF HUMAN OR ANIMAL CELLS
  JOURNAL   Patent: EP 0576862-A 3 05-JAN-1994;
            BOEHRINGER MANNHEIM GMBH (DE)
FEATURES             Location/Qualifiers
     source          1. .25
                     /organism="unidentified"
                     /db_xref="taxon:32644"
BASE COUNT        3 a      4 c      7 g     11 t
ORIGIN

  Query Match            66.0%;  Score 13.2;  DB 6;  Length 25;
  Best Local Similarity  83.3%;  Pred. No. 2.3e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 GATCTTGAACTTCCTCAT 19
          ||||||||  |||||| |
Db      1 GATCTTGAGTTTCCTCGT 18


RESULT 15
I68074
LOCUS       I68074                    25 bp    DNA     linear   PAT 04-FEB-1998
DEFINITION  Sequence 3 from patent US 5674723.
ACCESSION   I68074
VERSION     I68074.1  GI:2830196
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 25)
  AUTHORS   Abken,H.Johann., Willecke,K., Jungfer,H. and Barchet,H.
  TITLE     Nucleic acid molecules which immortalize human or animal cells and
            uses thereof
  JOURNAL   Patent: US 5674723-A 3 07-OCT-1997;
FEATURES             Location/Qualifiers
     source          1. .25
                     /organism="unknown"
BASE COUNT        3 a      4 c      7 g     11 t
ORIGIN

  Query Match            66.0%;  Score 13.2;  DB 6;  Length 25;
  Best Local Similarity  83.3%;  Pred. No. 2.3e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 GATCTTGAACTTCCTCAT 19
          ||||||||  |||||| |
Db      1 GATCTTGAGTTTCCTCGT 18


Search completed: December  3, 2002, 22:23:27
Job time : 358.3 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 06:29:45 ; Search time 94.1 Seconds
(without alignments)
478.639 Million cell updates/sec

Title: US-09-296-264-25
Perfect score: 20
Sequence: 1 cgatcttgaacttcctcatg 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 2390332

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	20	100.0	20	21	AAZ31455	Human neuropilin m
c	2	14.8	74.0	60	24	ABN44303	Human spliced tran
	3	14.2	71.0	34	22	AAF55693	Primer used to amp
c	4	14.2	71.0	65	24	ABN55280	Mouse spliced tran
	5	14.2	71.0	92	22	ABA73208	Human foetal liver
	6	14.2	71.0	92	22	ABA38641	Probe #17107 for g
	7	14.2	71.0	92	22	AAK21643	Human brain expres
	8	14.2	71.0	92	22	AAK47808	Human bone marrow
	9	14.2	71.0	92	22	AAI26044	Probe #15977 for g
	10	14.2	71.0	92	22	AAI53636	Probe #22322 used
	11	14.2	71.0	92	24	ABS21815	Human genome-deriv
	12	14	70.0	51	22	AAL27757	Human SNP oligonuc
	13	14	70.0	92	22	AAK16533	Human brain expres
	14	14	70.0	92	22	AAK42289	Human bone marrow
	15	14	70.0	92	22	AAI23059	Probe #12992 for g
	16	14	70.0	92	22	AAI48365	Probe #17051 used
	17	14	70.0	92	24	ABS16340	Human genome-deriv
	18	13.8	69.0	21	24	AAL47974	Human papillomavir
	19	13.8	69.0	31	21	AAA14713	Probe for cDNA enc
	20	13.8	69.0	60	20	AAX16855	Bovine casein gene
c	21	13.8	69.0	60	20	AAX16855	Bovine casein gene
	22	13.8	69.0	65	24	ABN55852	Mouse spliced tran
c	23	13.8	69.0	65	24	ABN57973	Mouse spliced tran
	24	13.6	68.0	32	22	AAS00476	Mouse FIRE EGF dom
c	25	13.6	68.0	46	22	AAD16898	Soybean oxidosqual
c	26	13.6	68.0	52	22	AAC85538	Antisense strand o
c	27	13.6	68.0	60	24	ABN32669	Human spliced tran
	28	13.6	68.0	60	24	ABN37759	Human spliced tran
	29	13.4	67.0	20	21	AAZ56922	ETS-1 gene specifi
c	30	13.4	67.0	27	22	AAF80427	PCR primer for cDN
	31	13.4	67.0	35	20	AAX55680	Truncated cellulas
	32	13.4	67.0	45	22	AAC82216	Human retrovirus D
c	33	13.4	67.0	58	18	AAV77680	Staphylococcus aur
c	34	13.2	66.0	20	21	AAZ93856	Primer for amplify
	35	13.2	66.0	24	22	AAF55651	Primer for human e
	36	13.2	66.0	25	14	AAQ53325	Oligonucleotide st
c	37	13.2	66.0	26	21	AAZ88835	Human heparanase P
c	38	13.2	66.0	31	22	AAI30228	Human single nucle
c	39	13.2	66.0	33	24	ABL41347	Human STTA protein
	40	13.2	66.0	38	22	AAF29710	Murine factor VII
c	41	13.2	66.0	41	24	ABL41349	Human STTA protein
c	42	13.2	66.0	41	24	ABL41350	Human STTA protein
c	43	13.2	66.0	42	14	AAQ37897	Beta-casein probe
c	44	13.2	66.0	42	14	AAQ38099	Sequence of probe
c	45	13.2	66.0	42	19	AAV25595	Probe for human be

ALIGNMENTS

RESULT 1
AAZ31455
ID AAZ31455 standard; DNA; 20 BP.
XX
AC AAZ31455;
XX
DT 07-FEB-2000 (first entry)
XX
DE Human neuropilin mRNA specific antisense oligo GTI3626.
XX
KW Neuropilin; human; growth; metastasis; tumor; neovascularisation;
KW cancer; papilloma; diabetic retinopathy; antisense; ss.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9955855-A2.
XX
PD 04-NOV-1999.
XX
PF 23-APR-1999; 99WO-CA00324.
XX
PR 23-APR-1998; 98US-0082791.
XX
PA (GENE-) GENESENSE TECHNOLOGIES INC.
XX
PI Wright JA, Young AH, Lee YS;
XX
DR WPI; 2000-023357/02.
XX
PT Antisense oligonucleotides that inhibit neuropilin expression, useful
PT for treating cancer -

XX
PS Claim 4; Page 17; 57pp; English.
XX
CC Sequences AAZ31431-460 represent antisense oligonucleotides which
CC inhibit human neuropilin expression. The antisense oligonucleotides can
CC be used to inhibit the growth or metastasis of a mammalian tumor and
CC inhibit neovascularisation. The oligonucleotides may be used to treat
CC various forms of cancers or tumors, such as sarcomas, melanomas,
CC adenomas, carcinomas of solid tissue, hypoxic tumors, squamous cell
CC carcinomas of the mouth, throat, larynx and lung, genitourinary cancers
CC such as cervical and bladder cancer, hematopoietic cancers, colon cancer,
CC breast cancer, pancreatic cancer, renal cancer, brain cancer, skin
CC cancer, liver cancer, head and neck cancers, and nervous system cancers,
CC as well as benign lesions such as papillomas. The methods may be used to
CC treat neovascularisation disorders such as diabetic retinopathy, and
CC retinopathy of prematurity and age related macular degeneration.
XX
SQ Sequence 20 BP; 4 A; 6 C; 3 G; 7 T; 0 other;

Query Match 100.0%; Score 20; DB 21; Length 20;
Best Local Similarity 100.0%; Pred. No. 3.1;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGATCTTGAACTTCCTCATG 20
||||||||||||||||||||
Db 1 CGATCTTGAACTTCCTCATG 20

RESULT 2
ABN44303/c
ID ABN44303 standard; DNA; 60 BP.
XX
AC ABN44303;
XX
DT 15-JUL-2002 (first entry)
XX
DE Human spliced transcript detection oligonucleotide SEQ ID NO:17051.
XX
KW Human; mouse; rat; splice transcript; detection; RNA transcript;
KW splice variant; transcriptome; oligonucleotide library; ss.
XX
OS Homo sapiens.
XX
PN WO200210449-A2.
XX
PD 07-FEB-2002.
XX
PF 20-JUL-2001; 2001WO-IB01903.
XX
PR 28-JUL-2000; 2000US-221607P.
PR 02-MAY-2001; 2001US-287724P.
XX
PA (COMP-) COMPUGEN INC.
XX
PI Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
XX
DR WPI; 2002-257383/30.
XX
PT New oligonucleotide libraries comprising oligonucleotides which
PT selectively hybridize to mRNAs transcribed from a transcription unit of
PT a genome, useful for detecting tissue-, pathology-, and
PT developmental-specific genes -
XX
PS Example 1; SEQ ID 17051; 47pp; English.
XX
CC The present invention describes oligonucleotide libraries for detecting
CC messenger RNAs that populate a (sub-)transcriptome, where the
CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC transcription units that populate a genome. The library comprises
CC several oligonucleotides, each capable of hybridising selectively to a
CC set of messenger RNAs transcribed from a given transcription unit of
CC the genome, which encodes one or more messenger RNA splice variants.
CC The oligonucleotide libraries are useful for detecting mRNAs from a
CC biological sample, in expression profiling studies, in qualitatively or
CC quantitatively characterising the corresponding transcriptome, and in
CC detecting RNA transcripts and splice variants of human or animal
CC transcriptomes. The libraries may also be used as specialised mini
CC libraries to detect transcripts of a sub-transcriptome under a
CC particular biological or pathological state, and so allowing the
CC detection of tissue- and pathology-specific genes such as those genes
CC only expressed in specific tissue under a specific pathological
CC condition; to detect developmental specific genes; and to detect RNA
CC transcripts and splice variants of a transcriptome of a patient suffering
CC from a particular disorder. ABN27253 to ABN59589 represent
CC oligonucleotide sequences from rats, humans and mice, which are used in
CC the exemplification of the present invention.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 60 BP; 26 A; 12 C; 9 G; 13 T; 0 other;

Query Match 74.0%; Score 14.8; DB 24; Length 60;
Best Local Similarity 88.9%; Pred. No. 1.1e+03;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 GATCTTGAACTTCCTCAT 19
| ||||||| ||||||||
Db 30 GCTCTTGAATTTCCTCAT 13

RESULT 3
AAF55693
ID AAF55693 standard; DNA; 34 BP.
XX
AC AAF55693;
XX
DT 11-JUN-2001 (first entry)
XX
DE Primer used to amplify the hemEH genes of Bacillus subtilis.
XX
KW hemE gene; uroporphyrinogen decarboxylase; hemH gene; ferrochelatase;
KW protohaem IX; leaven; lactic acid bacteria; respiratory metabolism;
KW fermented product; PCR primer; ss.
XX
OS Bacillus subtilis.
XX
PN WO200121808-A2.
XX
PD 29-MAR-2001.
XX
PF 20-SEP-2000; 2000WO-FR02611.
XX
PR 20-SEP-1999; 99FR-0011735.
XX
PA (INRG) INRA INST NAT RECH AGRONOMIQUE.
PA (DUWA/) DUWAT C.
PA (DUWA/) DUWAT C.
XX
PI Duwat P, Gruss A, Le Loir Y, Gaudu P;
XX
DR WPI; 2001-257993/26.
XX
PT Recombinant lactic acid bacterium with respiratory metabolism, useful
PT for producing fermented foods or recombinant proteins -
XX
PS Example 1; Page 18; 30pp; French.
XX
CC PCR primers AAF55692-93 were used to amplify hemE (uroporphyrinogen
CC decarboxylase) and hemH (ferrochelatase) genes. These genes are involved
CC in the synthesis of protohaem IX from glutamyl-tRNA. The amplified
CC sequence was used to produce recombinant lactic acid bacteria of the
CC invention. The specification describes recombinant lactic acid bacteria
CC which have been genetically modified to confer, or activate, a
CC respiratory metabolism. Lactic bacteria with a respiratory metabolism
CC show improved yield and greater viability during storage (comparable to

```
CC   the effect achieved by adding haem or other porphyrins). Cultures of
CC   the recombinant bacteria are useful as leavens for preparation of
CC   fermented products and for recombinant production of heterologous
CC   proteins.
XX
SQ   Sequence 34 BP; 9 A; 8 C; 2 G; 15 T; 0 other;

  Query Match            71.0%;  Score 14.2;  DB 22;  Length 34;
  Best Local Similarity  84.2%;  Pred. No. 2e+03;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 CGATCTTGAACTTCCTCAT 19
          ||||||| ||| |||| ||
Db      5 CGATCTTTAACGTCCTAAT 23


RESULT 4
ABN55280/c
ID   ABN55280 standard; DNA; 65 BP.
XX
AC   ABN55280;
XX
DT   15-JUL-2002  (first entry)
XX
DE   Mouse spliced transcript detection oligonucleotide SEQ ID NO:28028.
XX
KW   Human; mouse; rat; splice transcript; detection; RNA transcript;
KW   splice variant; transcriptome; oligonucleotide library; ss.
XX
OS   Mus musculus.
XX
PN   WO200210449-A2.
XX
PD   07-FEB-2002.
XX
PF   20-JUL-2001; 2001WO-IB01903.
XX
PR   28-JUL-2000; 2000US-221607P.
PR   02-MAY-2001; 2001US-287724P.
XX
PA   (COMP-) COMPUGEN INC.
XX
PI   Shoshan A,  Wasserman A,  Mintz E,  Mintz L,  Faigler S;
XX
DR   WPI; 2002-257383/30.
XX
PT   New oligonucleotide libraries comprising oligonucleotides which
PT   selectively hybridize to mRNAs transcribed from a transcription unit of
PT   a genome, useful for detecting tissue-, pathology-, and
PT   developmental-specific genes  -
XX
PS   Example 1; SEQ ID 28028; 47pp; English.
XX
CC   The present invention describes oligonucleotide libraries for detecting
CC   messenger RNAs that populate a (sub-)transcriptome, where the
CC   (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC   transcription units that populate a genome. The library comprises
CC   several oligonucleotides, each capable of hybridising selectively to a
CC   set of messenger RNAs transcribed from a given transcription unit of
CC   the genome, which encodes one or more messenger RNA splice variants.
CC   The oligonucleotide libraries are useful for detecting mRNAs from a
CC   biological sample, in expression profiling studies, in qualitatively or
CC   quantitatively characterising the corresponding transcriptome, and in
CC   detecting RNA transcripts and splice variants of human or animal
CC   transcriptomes. The libraries may also be used as specialised mini
CC   libraries to detect transcripts of a sub-transcriptome under a
CC   particular biological or pathological state, and so allowing the
CC   detection of tissue- and pathology-specific genes such as those genes
CC   only expressed in specific tissue under a specific pathological
CC   condition; to detect developmental specific genes; and to detect RNA
CC   transcripts and splice variants of a transcriptome of a patient suffering
CC   from a particular disorder. ABN27253 to ABN59589 represent
CC   oligonucleotide sequences from rats, humans and mice, which are used in
CC   the exemplification of the present invention.
CC   N.B. The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 65 BP; 22 A; 16 C; 14 G; 13 T; 0 other;

  Query Match            71.0%;  Score 14.2;  DB 24;  Length 65;
  Best Local Similarity  84.2%;  Pred. No. 2.2e+03;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 GATCTTGAACTTCCTCATG 20
          || ||||| |||||||| |
Db     25 GAACTTGACCTTCCTCAAG 7


RESULT 5
ABA73208
ID   ABA73208 standard; DNA; 92 BP.
XX
AC   ABA73208;
XX
DT   01-FEB-2002  (first entry)
XX
DE   Human foetal liver single exon nucleic acid probe #21513.
XX
KW   Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157277-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00669.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-483447/52.
XX
PT   Human genome-derived single exon nucleic acid probes useful for
PT   analyzing gene expression in human fetal liver -
XX
PS   Claim 4; SEQ ID NO 21513; 639pp + sequence listing; English.
XX
CC   The invention relates to a single exon nucleic acid probe for
CC   measuring human gene expression in a sample derived from human foetal
CC   liver. The single exon nucleic acid probes may be used for predicting,
CC   measuring and displaying gene expression in samples derived from human
CC   fetal liver. The present sequence is a single exon nucleic acid
CC   probe of the invention.
CC   Note: The sequence data for this patent did not form part of the
CC   printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 92 BP; 17 A; 29 C; 11 G; 35 T; 0 other;

  Query Match            71.0%;  Score 14.2;  DB 22;  Length 92;
  Best Local Similarity  84.2%;  Pred. No. 2.3e+03;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 GATCTTGAACTTCCTCATG 20
          |||||||||  ||| ||||
```

```
Db     29 GATCTTGAAAATCCCCATG 47


RESULT 6
ABA38641
ID   ABA38641 standard; DNA; 92 BP.
XX
AC   ABA38641;
XX
DT   23-JAN-2002  (first entry)
XX
DE   Probe #17107 for gene expression analysis in human heart cell sample.
XX
KW   Human; gene expression; heart; microarray; vascular system; probe;
KW   cardiovascular disease; hypertension; cardiac arrhythmia;
KW   congenital heart disease; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157274-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00666.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488899/53.
XX
PT   Single exon nucleic acid probes for analyzing gene expression in human
PT   hearts -
XX
PS   Claim 4; SEQ ID No 17107; 530pp; English.
XX
CC   The present invention relates to single exon nucleic acid probes for
CC   measuring human gene expression in a sample derived from human heart. The
CC   present sequence is one such probe. The probes may be used for
CC   predicting, measuring and displaying gene expression in samples derived
CC   from the human heart via microarrays. By measuring gene expression, the
CC   probes are useful for predicting, diagnosing, grading, staging,
CC   monitoring and prognosing diseases of the human heart and vascular system
CC   e.g. cardiovascular disease, hypertension, cardiac arrhythmias and
CC   congenital heart disease.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 92 BP; 17 A; 29 C; 11 G; 35 T; 0 other;

  Query Match             71.0%;  Score 14.2;  DB 22;  Length 92;
  Best Local Similarity   84.2%;  Pred. No. 2.3e+03;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 GATCTTGAACTTCCTCATG 20
           |||||||||  ||| ||||
Db      29 GATCTTGAAAATCCCCATG 47


RESULT 7
AAK21643
ID   AAK21643 standard; DNA; 92 BP.
XX
AC   AAK21643;
XX
DT   05-NOV-2001  (first entry)
XX
DE   Human brain expressed single exon probe SEQ ID NO: 21634.
XX
KW   Human; brain expressed exon; gene expression analysis; probe;
KW   microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
KW   epilepsy; cancer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157275-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00667.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-483446/52.
XX
PT   Single exon nucleic acid probes for analyzing gene expression in human
PT   brains -
XX
PS   Example 4; SEQ ID NO: 21634; 650pp + Sequence Listing; English.
XX
CC   The present invention provides a number of single exon nucleic acid
CC   probes which are derived from genomic sequences expressed in the human
CC   brain. They can be used to measure gene expression in brain cell samples,
CC   which may enable the diagnosis and improved treatment of nervous system
CC   diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC   epilepsy and cancers. The present sequence is one of the probes of the
CC   invention.
XX
SQ   Sequence 92 BP; 17 A; 29 C; 11 G; 35 T; 0 other;

  Query Match             71.0%;  Score 14.2;  DB 22;  Length 92;
  Best Local Similarity   84.2%;  Pred. No. 2.3e+03;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 GATCTTGAACTTCCTCATG 20
           |||||||||  ||| ||||
Db      29 GATCTTGAAAATCCCCATG 47


RESULT 8
AAK47808
ID   AAK47808 standard; DNA; 92 BP.
XX
AC   AAK47808;
XX
DT   06-NOV-2001  (first entry)
XX
DE   Human bone marrow expressed single exon probe SEQ ID NO: 22365.
XX
KW   Human; bone marrow expressed exon; gene expression analysis; probe;
KW   microarray; cancer; leukaemia; lymphoma; myeloma; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157276-A2.
XX
PD   09-AUG-2001.
```

```
XX
PF   30-JAN-2001; 2001WO-US00668.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488900/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for
PT   analyzing gene expression in human bone marrow -
XX
PS   Example 4; SEQ ID NO: 22365; 658pp + Sequence Listing; English.
XX
CC   The present invention provides a number of single exon nucleic acid
CC   probes which are derived from genomic sequences expressed in the human
CC   bone marrow. They can be used to measure gene expression in bone marrow
CC   samples, which may enable the improved diagnosis and treatment of cancers
CC   such as lymphoma, leukaemia and myeloma. The present sequence is one of
CC   the probes of the invention.
XX
SQ   Sequence 92 BP; 17 A; 29 C; 11 G; 35 T; 0 other;

  Query Match             71.0%;  Score 14.2;  DB 22;  Length 92;
  Best Local Similarity   84.2%;  Pred. No. 2.3e+03;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 GATCTTGAACTTCCTCATG 20
           |||||||||  ||| ||||
Db      29 GATCTTGAAAATCCCCATG 47


RESULT 9
AAI26044
ID   AAI26044 standard; DNA; 92 BP.
XX
AC   AAI26044;
XX
DT   12-OCT-2001  (first entry)
XX
DE   Probe #15977 for gene expression analysis in human cervical cell sample.
XX
KW   Probe; human; microarray; gene expression; cervical epithelial cell;
KW   cervical cancer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157278-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00670.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488901/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for
PT   analyzing gene expression in human cervical epithelial cells -
XX
PS   Claim 25; SEQ ID No 15977; 487pp; English.
XX
CC   The present invention relates to human single exon nucleic acid probes
CC   (SENP). The present sequence is one such probe. The SENPs are derived
CC   from human HeLa cells. The SENPs can be used to produce a single exon
CC   microarray, which can be used for measuring human gene expression in a
CC   sample derived from human cervical epithelial cells. By measuring gene
CC   expression, the probes are therefore useful in grading and/or staging
CC   of diseases of the cervix, notably cervical cancer.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 92 BP; 17 A; 29 C; 11 G; 35 T; 0 other;

  Query Match             71.0%;  Score 14.2;  DB 22;  Length 92;
  Best Local Similarity   84.2%;  Pred. No. 2.3e+03;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 GATCTTGAACTTCCTCATG 20
           |||||||||  ||| ||||
Db      29 GATCTTGAAAATCCCCATG 47


RESULT 10
AAI53636
ID   AAI53636 standard; DNA; 92 BP.
XX
AC   AAI53636;
XX
DT   17-OCT-2001  (first entry)
XX
DE   Probe #22322 used to measure gene expression in human placenta sample.
XX
KW   Probe; microarray; human; placenta; antenatal diagnosis;
KW   genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157272-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00663.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488897/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for
PT   analyzing gene expression in human placenta -
XX
PS   Claim 25; SEQ ID No 22322; 654pp; English.
XX
CC   The present invention relates to single exon nucleic acid probes (SENP).
CC   The present sequence is one such probe. The probes are useful for
CC   producing a microarray for predicting, measuring and displaying gene
CC   expression in samples derived from human placenta. The probes are useful
```

```
CC   for antenatal diagnosis of human genetic disorders.
XX
SQ   Sequence 92 BP; 17 A; 29 C; 11 G; 35 T; 0 other;

  Query Match             71.0%;  Score 14.2;  DB 22;  Length 92;
  Best Local Similarity   84.2%;  Pred. No. 2.3e+03;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 GATCTTGAACTTCCTCATG 20
          |||||||||  ||| ||||
Db     29 GATCTTGAAAATCCCCATG 47


RESULT 11
ABS21815
ID   ABS21815 standard; DNA; 92 BP.
XX
AC   ABS21815;
XX
DT   19-AUG-2002  (first entry)
XX
DE   Human genome-derived single exon probe ORF from lung SEQ ID No 21806.
XX
KW   Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;
KW   chronic obstructive pulmonary disease; interstitial lung disease;
KW   familial idiopathic pulmonary fibrosis; neurofibromatosis;
KW   tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
KW   Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
KW   pulmonary histiocytosis; lymphangioleiomyomtosis; Karagener syndrome;
KW   pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
KW   primary ciliary dyskinesis; pulmonary hypertension;
KW   hyaline membrane disease; open reading frame; ORF.
XX
OS   Homo sapiens.
XX
PN   WO200186003-A2.
XX
PD   15-NOV-2001.
XX
PF   30-JAN-2001; 2001WO-US00665.
XX
PR   04-FEB-2000; 2000US-180312P.
PR   26-MAY-2000; 2000US-207456P.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-234687P.
PR   27-SEP-2000; 2000US-236359P.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2002-114183/15.
XX
PT   Spatially-addressable set of single exon nucleic acid probes, used to
PT   measure gene expression in human lung samples -
XX
PS   Claim 4; SEQ ID No 21806; 634pp; English.
XX
CC   The invention relates to a spatially-addressable set of single exon
CC   nucleic acid probes for measuring gene expression in a sample derived
CC   from human lung comprising single exon nucleic acid probes having one of
CC   12614 nucleic acid sequences mentioned in the specification, or their
CC   complements or the 12387 open reading frames derived from the 12614
CC   probes. Also included are a microarray comprising the novel set of
CC   probes; the novel set of probes which hybridise at high stringency to a
CC   nucleic acid expressed in the human lung; measuring gene expression in a
CC   sample derived from human lung, comprising (a) contacting the array with
CC   a collection of detectably labeled nucleic acids derived from human lung
CC   mRNA, and (b) measuring the label detectably bound to each probe of
CC   the array; identifying exons in a eukaryotic genome, comprising
CC   (a) algorithmically predicting at least one exon from genomic sequences
CC   of the eukaryote; and (b) detecting specific hybridisation of detectably
CC   labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
CC   having a fragment identical to the predicted exon, the probe is included
CC   in the above mentioned microarray; assigning exons to a single gene,
CC   comprising (a) identifying exons from genomic sequence by the method
CC   above and (b) measuring the expression of each of the exons in several
CC   tissues and/or cell types using hybridisation to a single exon
CC   microarrays having a probe with the exon, where a common pattern of
CC   expression of the exons in the tissues and/or cell types indicates that
CC   the exons should be assigned to a single gene; a peptide comprising one
CC   of 12011 sequences, mentioned in the specification, or encoded by the
CC   probes/open reading frames (ORF). The probes are used for gene
CC   expression analysis, and for identifying exons in a gene, particularly
CC   using human lung derived mRNA and for the study of lung diseases
CC   such as asthma, lung cancer, chronic obstructive pulmonary disease
CC   (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
CC   fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,
CC   Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary
CC   haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomtosis,
CC   pulmonary alveolar proteinosis, Karagener syndrome, fibrocystic
CC   pulmonary dysplasia, primary ciliary dyskinesis, pulmonary hypertension
CC   and hyaline membrane disease. The present sequence is a single exon
CC   probe open reading frame of the invention.
CC   Note: The sequence data for this patent did not form part
CC   of the printed specification, but was obtained in electronic
CC   format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 92 BP; 17 A; 29 C; 11 G; 35 T; 0 other;

  Query Match             71.0%;  Score 14.2;  DB 24;  Length 92;
  Best Local Similarity   84.2%;  Pred. No. 2.3e+03;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 GATCTTGAACTTCCTCATG 20
          |||||||||  ||| ||||
Db     29 GATCTTGAAAATCCCCATG 47


RESULT 12
AAL27757
ID   AAL27757 standard; DNA; 51 BP.
XX
AC   AAL27757;
XX
DT   24-JAN-2002  (first entry)
XX
DE   Human SNP oligonucleotide #965.
XX
KW   Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
KW   neuroprotective; antimicrobial; gene therapy; vaccine; amylase; cancer;
KW   amyloid protein; angiopoietin; apoptosis related protein; cadherin;
KW   cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
KW   complement related protein; cytochrome; kinesin; cytokine; interferon;
KW   interleukin; G-protein coupled receptor; thioesterase; inflammation;
KW   multifactorial disease; autoimmune disease; infection;
KW   nervous system disease; ss.
XX
OS   Homo sapiens.
XX
PN   WO200147944-A2.
XX
PD   05-JUL-2001.
XX
PF   28-DEC-2000; 2000WO-US35498.
XX
PR   28-DEC-1999;   99US-0173419.
PR   27-DEC-2000; 2000US-0173419.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Shimkets RA,  Leach M;
XX
```

```
DR   WPI; 2001-465210/50.
XX
PT   Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
PT   oncogenes and histones, useful for diagnosing and treating, e.g.
PT   cancer, autoimmune diseases and infections -
XX
PS   Claim 1; Page 1656; 4143pp; English.
XX
CC   The present invention relates to oligonucleotides encoding polymorphic
CC   variants of proteins related to amylases, amyloid proteins, angiopoietin,
CC   apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,
CC   histones, kinases, colony stimulating factors, complement related
CC   proteins, cytochromes, kinesins, cytokines, interferons, interleukins,
CC   G-protein coupled receptors and thioesterases. The present sequence is
CC   one such oligonucleotide. The oligonucleotides and the peptides encoded
CC   by them may be used in the prevention, diagnosis and treatment of
CC   diseases associated with inappropriate expression of the proteins listed
CC   above. Disorders that may be prevented, diagnosed and/or treated include
CC   multifactorial diseases with a genetic component, such as autoimmune
CC   diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,
CC   systemic lupus erythromatosus and Grave's disease), inflammation, cancer
CC   (e.g. cancers of the bladder, brain, breast, colon and kidney,
CC   leukaemia), diseases of the nervous system and an infection of pathogenic
CC   organisms.
XX
SQ   Sequence 51 BP; 11 A; 13 C; 11 G; 16 T; 0 other;

  Query Match             70.0%;  Score 14;  DB 22;  Length 51;
  Best Local Similarity   100.0%;  Pred. No. 2.6e+03;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CGATCTTGAACTTC 14
           ||||||||||||||
Db      12 CGATCTTGAACTTC 25


RESULT 13
AAK16533
ID   AAK16533 standard; DNA; 92 BP.
XX
AC   AAK16533;
XX
DT   05-NOV-2001  (first entry)
XX
DE   Human brain expressed single exon probe SEQ ID NO: 16524.
XX
KW   Human; brain expressed exon; gene expression analysis; probe;
KW   microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
KW   epilepsy; cancer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157275-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00667.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-483446/52.
XX
PT   Single exon nucleic acid probes for analyzing gene expression in human
PT   brains -
XX
PS   Example 4; SEQ ID NO: 16524; 650pp + Sequence Listing; English.
XX
CC   The present invention provides a number of single exon nucleic acid
CC   probes which are derived from genomic sequences expressed in the human
CC   brain. They can be used to measure gene expression in brain cell samples,
CC   which may enable the diagnosis and improved treatment of nervous system
CC   diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC   epilepsy and cancers. The present sequence is one of the probes of the
CC   invention.
XX
SQ   Sequence 92 BP; 27 A; 27 C; 14 G; 24 T; 0 other;

  Query Match             70.0%;  Score 14;  DB 22;  Length 92;
  Best Local Similarity   100.0%;  Pred. No. 2.8e+03;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 GATCTTGAACTTCC 15
           ||||||||||||||
Db      43 GATCTTGAACTTCC 56


RESULT 14
AAK42289
ID   AAK42289 standard; DNA; 92 BP.
XX
AC   AAK42289;
XX
DT   06-NOV-2001  (first entry)
XX
DE   Human bone marrow expressed single exon probe SEQ ID NO: 16846.
XX
KW   Human; bone marrow expressed exon; gene expression analysis; probe;
KW   microarray; cancer; leukaemia; lymphoma; myeloma; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157276-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00668.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488900/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for
PT   analyzing gene expression in human bone marrow -
XX
PS   Example 4; SEQ ID NO: 16846; 658pp + Sequence Listing; English.
XX
CC   The present invention provides a number of single exon nucleic acid
CC   probes which are derived from genomic sequences expressed in the human
CC   bone marrow. They can be used to measure gene expression in bone marrow
CC   samples, which may enable the improved diagnosis and treatment of cancers
CC   such as lymphoma, leukaemia and myeloma. The present sequence is one of
CC   the probes of the invention.
XX
SQ   Sequence 92 BP; 27 A; 27 C; 14 G; 24 T; 0 other;

  Query Match             70.0%;  Score 14;  DB 22;  Length 92;
```

```
  Best Local Similarity   100.0%;  Pred. No. 2.8e+03;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 GATCTTGAACTTCC 15
           ||||||||||||||
Db      43 GATCTTGAACTTCC 56


RESULT 15
AAI23059
ID   AAI23059 standard; DNA; 92 BP.
XX
AC   AAI23059;
XX
DT   12-OCT-2001  (first entry)
XX
DE   Probe #12992 for gene expression analysis in human cervical cell sample.
XX
KW   Probe; human; microarray; gene expression; cervical epithelial cell;
KW   cervical cancer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157278-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00670.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488901/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for
PT   analyzing gene expression in human cervical epithelial cells -
XX
PS   Claim 25; SEQ ID No 12992; 487pp; English.
XX
CC   The present invention relates to human single exon nucleic acid probes
CC   (SENP). The present sequence is one such probe. The SENPs are derived
CC   from human HeLa cells. The SENPs can be used to produce a single exon
CC   microarray, which can be used for measuring human gene expression in a
CC   sample derived from human cervical epithelial cells. By measuring gene
CC   expression, the probes are therefore useful in grading and/or staging
CC   of diseases of the cervix, notably cervical cancer.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 92 BP; 27 A; 27 C; 14 G; 24 T; 0 other;

  Query Match             70.0%;  Score 14;  DB 22;  Length 92;
  Best Local Similarity   100.0%;  Pred. No. 2.8e+03;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       2 GATCTTGAACTTCC 15
           ||||||||||||||
Db      43 GATCTTGAACTTCC 56


Search completed: November 23, 2002, 07:03:51
Job time : 97.1 secs
```

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 06:42:25 ; Search time 16.8 Seconds
(without alignments)
450.869 Million cell updates/sec

Title: US-09-296-264-25
Perfect score: 20
Sequence: 1 cgatcttgaacttcctcatg 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 335578 seqs, 189365133 residues

Total number of hits satisfying chosen parameters: 195614

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
c	1	15.2	76.0	98	10	US-09-878-574-12594	Sequence 12594, A
	2	14.2	71.0	92	10	US-09-864-761-23961	Sequence 23961, A
c	3	14.2	71.0	98	10	US-09-878-574-11296	Sequence 11296, A
	4	14	70.0	92	10	US-09-864-761-26932	Sequence 26932, A
c	5	13.8	69.0	67	9	US-10-079-623-312	Sequence 312, App
c	6	13.6	68.0	52	10	US-09-808-037-14	Sequence 14, Appl
c	7	13.4	67.0	27	9	US-10-010-928-5	Sequence 5, Appli
c	8	13.4	67.0	27	10	US-09-802-124-5	Sequence 5, Appli
c	9	13.4	67.0	27	10	US-09-733-605-5	Sequence 5, Appli
c	10	13.2	66.0	85	10	US-09-878-574-5344	Sequence 5344, Ap
c	11	13.2	66.0	91	10	US-09-864-761-21961	Sequence 21961, A
	12	13.2	66.0	96	10	US-09-878-574-13097	Sequence 13097, A
	13	13	65.0	21	10	US-09-231-235-39	Sequence 39, Appl
	14	13	65.0	21	10	US-09-797-518A-39	Sequence 39, Appl
	15	13	65.0	33	10	US-09-231-235-9	Sequence 9, Appli
	16	13	65.0	33	10	US-09-797-518A-9	Sequence 9, Appli
c	17	12.8	64.0	31	10	US-09-801-274-716	Sequence 716, App
c	18	12.8	64.0	47	9	US-09-853-526-330	Sequence 330, App
c	19	12.8	64.0	47	10	US-09-901-484A-330	Sequence 330, App
c	20	12.8	64.0	80	9	US-09-970-616-4	Sequence 4, Appli
c	21	12.6	63.0	22	10	US-09-804-386-1	Sequence 1, Appli
	22	12.6	63.0	93	10	US-09-960-352-1166	Sequence 1166, Ap
c	23	12.6	63.0	93	10	US-09-960-352-4510	Sequence 4510, Ap
c	24	12.6	63.0	95	10	US-09-920-300A-1545	Sequence 1545, Ap
c	25	12.6	63.0	95	12	US-10-033-528-1545	Sequence 1545, Ap
	26	12.4	62.0	96	10	US-09-864-761-18275	Sequence 18275, A
	27	12.4	62.0	99	10	US-09-864-761-17487	Sequence 17487, A
c	28	12.4	62.0	99	10	US-09-864-761-28969	Sequence 28969, A
	29	12.2	61.0	23	10	US-09-966-147-31	Sequence 31, Appl
	30	12.2	61.0	24	10	US-09-935-785-2	Sequence 2, Appli
	31	12.2	61.0	31	10	US-09-875-338-91	Sequence 91, Appl
c	32	12.2	61.0	33	9	US-10-051-311A-3	Sequence 3, Appli
	33	12.2	61.0	45	9	US-09-925-922-7	Sequence 7, Appli
c	34	12.2	61.0	45	10	US-09-962-055-30	Sequence 30, Appl
c	35	12.2	61.0	45	12	US-10-023-529-30	Sequence 30, Appl
c	36	12.2	61.0	45	12	US-10-023-523-30	Sequence 30, Appl
c	37	12.2	61.0	78	10	US-09-294-093B-4129	Sequence 4129, Ap
c	38	12.2	61.0	78	10	US-09-962-055-31	Sequence 31, Appl
c	39	12.2	61.0	78	12	US-10-023-529-31	Sequence 31, Appl
c	40	12.2	61.0	78	12	US-10-023-523-31	Sequence 31, Appl
	41	12.2	61.0	99	10	US-09-815-242-1817	Sequence 1817, Ap
	42	12.2	61.0	99	10	US-09-815-242-2514	Sequence 2514, Ap
c	43	12.2	61.0	100	10	US-09-864-761-24515	Sequence 24515, A
c	44	12.2	61.0	100	10	US-09-864-761-31235	Sequence 31235, A
c	45	12	60.0	25	10	US-09-828-313-57	Sequence 57, Appl

ALIGNMENTS

RESULT 1
US-09-878-574-12594/c
; Sequence 12594, Application US/09878574
; Patent No. US20020110548A1
; GENERAL INFORMATION:
; APPLICANT: Byrum, Joseph R.
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Thompson, Michael D.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(15401)B
; CURRENT APPLICATION NUMBER: US/09/878,574
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 09/333,535
; PRIOR FILING DATE: 1999-06-14
; NUMBER OF SEQ ID NOS: 15775
; SEQ ID NO 12594
; LENGTH: 98
; TYPE: DNA
; ORGANISM: Glycine max
; OTHER INFORMATION: Clone ID: 701065965H1
US-09-878-574-12594

Query Match 76.0%; Score 15.2; DB 10; Length 98;
Best Local Similarity 85.0%; Pred. No. 1.1e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy       1 CGATCTTGAACTTCCTCATG 20
           |||| ||||||||| || ||
Db      60 CGATGTTGAACTTCATCTTG 41
```

RESULT 2
US-09-864-761-23961
; Sequence 23961, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO

```
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 23961
;   LENGTH: 92
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AL133419.11
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL =   4
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 4.3
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 4.3
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.3
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 4.4
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL =   4
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL =   4
;   OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 4.1
;   OTHER INFORMATION: NT HIT: AL161589.2, EVALUE 3.00e-03
US-09-864-761-23961

  Query Match             71.0%;  Score 14.2;  DB 10;  Length 92;
  Best Local Similarity   84.2%;  Pred. No. 3.3e+02;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 GATCTTGAACTTCCTCATG 20
          |||||||||  ||| ||||
Db     29 GATCTTGAAAATCCCCATG 47


RESULT 3
US-09-878-574-11296/c
; Sequence 11296, Application US/09878574
; Patent No. US20020110548A1
; GENERAL INFORMATION:
;  APPLICANT: Byrum, Joseph R.
;  APPLICANT:  La Rosa, Thomas J.
;  APPLICANT:  Thompson, Michael D.
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
;  TITLE OF INVENTION:  Plants
;  FILE REFERENCE: 38-21(15401)B
;  CURRENT APPLICATION NUMBER: US/09/878,574
;  CURRENT FILING DATE:  2001-12-21
;  PRIOR APPLICATION NUMBER: 09/333,535
;  PRIOR FILING DATE: 1999-06-14
;  NUMBER OF SEQ ID NOS: 15775
; SEQ ID NO 11296
;   LENGTH: 98
;   TYPE: DNA
;   ORGANISM: Glycine max
;   OTHER INFORMATION: Clone ID: 701064281H1
US-09-878-574-11296

  Query Match             71.0%;  Score 14.2;  DB 10;  Length 98;
  Best Local Similarity   84.2%;  Pred. No. 3.4e+02;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 CGATCTTGAACTTCCTCAT 19
          |||||||||  ||| | |||
Db     65 CGATCTTGACCTTGCCCAT 47


RESULT 4
US-09-864-761-26932
; Sequence 26932, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
```

```
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 26932
;   LENGTH: 92
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AC006317.3
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.63
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.61
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.68
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.61
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.76
;   OTHER INFORMATION: NT HIT: AL163209.2, EVALUE 3.00e-12
;   OTHER INFORMATION: SWISSPROT HIT: P13233, EVALUE 9.20e+00
;   OTHER INFORMATION: EST_HUMAN HIT: AA729888.1, EVALUE 6.00e-18
US-09-864-761-26932

  Query Match             70.0%;  Score 14;  DB 10;  Length 92;
  Best Local Similarity   100.0%;  Pred. No. 4.2e+02;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      2 GATCTTGAACTTCC 15
          ||||||||||||||
Db     43 GATCTTGAACTTCC 56


RESULT 5
US-10-079-623-312/c
; Sequence 312, Application US/10079623
; Patent No. US20020169302A1
; GENERAL INFORMATION:
;  APPLICANT: Havukkala, Ilkka J.
;  APPLICANT:  Glenn, Matthew
;  APPLICANT:  Grigor, Murray R.
;  APPLICANT:  Molenaar, Adrian J.
;  TITLE OF INVENTION: Compositions isolated from bovine
;  TITLE OF INVENTION:  mammary gland and methods for their use.
;  FILE REFERENCE: 11000.1044c3
;  CURRENT APPLICATION NUMBER: US/10/079,623
;  CURRENT FILING DATE:  2002-02-19
;  NUMBER OF SEQ ID NOS: 370
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 312
;   LENGTH: 67
;   TYPE: DNA
;   ORGANISM: Bovine
US-10-079-623-312

  Query Match             69.0%;  Score 13.8;  DB 9;  Length 67;
  Best Local Similarity   88.2%;  Pred. No. 5e+02;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      4 TCTTGAACTTCCTCATG 20
          ||| | ||||||||||
Db     46 TCTAGTACTTCCTCATG 30


RESULT 6
US-09-808-037-14/c
; Sequence 14, Application US/09808037
; Patent No. US20020052311A1
; GENERAL INFORMATION:
;  APPLICANT: SOLOMON, Beka
;  APPLICANT:  HANAN, Eilat
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE TREATMENT AND/OR DIAGNOSIS OF
;  TITLE OF INVENTION:  NEUROLOGICAL DISEASES AND DISORDERS
;  FILE REFERENCE: SOLOMON=2D
;  CURRENT APPLICATION NUMBER: US/09/808,037
;  CURRENT FILING DATE:  2001-03-15
;  PRIOR APPLICATION NUMBER: 09/629,971
;  PRIOR FILING DATE: 2000-07-31
;  PRIOR APPLICATION NUMBER: US 09/473,653
;  PRIOR FILING DATE: 1999-12-29
;  PRIOR APPLICATION NUMBER: US 60/152,417
;  PRIOR FILING DATE: 1999-09-03
;  NUMBER OF SEQ ID NOS: 33
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 14
;   LENGTH: 52
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: primer
US-09-808-037-14

  Query Match             68.0%;  Score 13.6;  DB 10;  Length 52;
  Best Local Similarity   80.0%;  Pred. No. 6e+02;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 CGATCTTGAACTTCCTCATG 20
          ||||  |||| ||| |||||
Db     51 CGATGCTGAATTTCGTCATG 32


RESULT 7
US-10-010-928-5/c
; Sequence 5, Application US/10010928
; Patent No. US20020155468A1
; GENERAL INFORMATION:
;  APPLICANT: Dillon, Davin C.
;  APPLICANT:  Hand-Zimmermann, Susan
;  APPLICANT:  Fling, Steven P.
;  TITLE OF INVENTION: OVARIAN TUMOR ANTIGEN AND METHODS OF USE
;  TITLE OF INVENTION:  THEREFOR
;  FILE REFERENCE: 210121.481C3
;  CURRENT APPLICATION NUMBER: US/10/010,928
;  CURRENT FILING DATE:  2001-12-07
;  NUMBER OF SEQ ID NOS: 39
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5
;   LENGTH: 27
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: HPP14 primer sequence
US-10-010-928-5

  Query Match             67.0%;  Score 13.4;  DB 9;  Length 27;
  Best Local Similarity   93.3%;  Pred. No. 6.7e+02;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      2 GATCTTGAACTTCCT 16
          ||||||||||||| |
Db     26 GATCTTGAACTTCTT 12


RESULT 8
US-09-802-124-5/c
; Sequence 5, Application US/09802124
; Patent No. US20020058292A1
; GENERAL INFORMATION:
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Zimmermann, Susan H.
;  APPLICANT:  Fling, Steven P.
;  TITLE OF INVENTION:  OVARIAN TUMOR ANTIGEN AND METHODS OF USE
;  TITLE OF INVENTION:  THEREFOR
;  FILE REFERENCE:  210121.481C2
;  CURRENT APPLICATION NUMBER: US/09/802,124
;  CURRENT FILING DATE:  2001-03-07
;  NUMBER OF SEQ ID NOS:  39
;  SOFTWARE:  FastSEQ for Windows Version 4.0
```

```
; SEQ ID NO 5
;   LENGTH: 27
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: HPP14 primer sequence
US-09-802-124-5

  Query Match             67.0%;  Score 13.4;  DB 10;  Length 27;
  Best Local Similarity   93.3%;  Pred. No. 6.7e+02;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      2 GATCTTGAACTTCCT 16
          ||||||||||||| |
Db     26 GATCTTGAACTTCTT 12


RESULT 9
US-09-733-605-5/c
; Sequence 5, Application US/09733605
; Patent No. US20020064815A1
; GENERAL INFORMATION:
;  APPLICANT:  Dillon, Davin C.
;  APPLICANT:  Zimmermann, Susan Hand
;  TITLE OF INVENTION:  OVARIAN TUMOR ANTIGEN AND METHODS OF USE
;  TITLE OF INVENTION:  THEREFOR
;  FILE REFERENCE:  210121.481C1
;  CURRENT APPLICATION NUMBER: US/09/733,605
;  CURRENT FILING DATE:  2000-12-08
;  NUMBER OF SEQ ID NOS:  39
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 5
;   LENGTH: 27
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: HPP14 primer sequence
US-09-733-605-5

  Query Match             67.0%;  Score 13.4;  DB 10;  Length 27;
  Best Local Similarity   93.3%;  Pred. No. 6.7e+02;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      2 GATCTTGAACTTCCT 16
          ||||||||||||| |
Db     26 GATCTTGAACTTCTT 12


RESULT 10
US-09-878-574-5344/c
; Sequence 5344, Application US/09878574
; Patent No. US20020110548A1
; GENERAL INFORMATION:
;  APPLICANT: Byrum, Joseph R.
;  APPLICANT:  La Rosa, Thomas J.
;  APPLICANT:  Thompson, Michael D.
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
;  TITLE OF INVENTION:  Plants
;  FILE REFERENCE: 38-21(15401)B
;  CURRENT APPLICATION NUMBER: US/09/878,574
;  CURRENT FILING DATE:  2001-12-21
;  PRIOR APPLICATION NUMBER: 09/333,535
;  PRIOR FILING DATE: 1999-06-14
;  NUMBER OF SEQ ID NOS: 15775
; SEQ ID NO 5344
;   LENGTH: 85
;   TYPE: DNA
;   ORGANISM: Glycine max
;   OTHER INFORMATION: Clone ID: LIB3028-050-Q1-B1-E11
US-09-878-574-5344

  Query Match             66.0%;  Score 13.2;  DB 10;  Length 85;
  Best Local Similarity   83.3%;  Pred. No. 1e+03;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 GATCTTGAACTTCCTCAT 19
          ||  ||| ||||||||||
Db     34 GACTTTGGACTTCCTCAT 17


RESULT 11
US-09-864-761-21961/c
; Sequence 21961, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 21961
;   LENGTH: 91
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AC004883.2
;   OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.5
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3.3
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.7
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.2
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.4
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL =   1
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.1
;   OTHER INFORMATION: NT HIT: AF038968.1, EVALUE 2.00e-44
```

```
;   OTHER INFORMATION: EST_HUMAN HIT: AL121131.1, EVALUE 3.00e-44
US-09-864-761-21961

  Query Match             66.0%;  Score 13.2;  DB 10;  Length 91;
  Best Local Similarity   83.3%;  Pred. No. 1e+03;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 CGATCTTGAACTTCCTCA 18
           |||  ||||||| |||||
Db      91 CGAAGTTGAACTCCCTCA 74


RESULT 12
US-09-878-574-13097
; Sequence 13097, Application US/09878574
; Patent No. US20020110548A1
; GENERAL INFORMATION:
;  APPLICANT: Byrum, Joseph R.
;  APPLICANT:  La Rosa, Thomas J.
;  APPLICANT:  Thompson, Michael D.
;  TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
;  TITLE OF INVENTION:  Plants
;  FILE REFERENCE: 38-21(15401)B
;  CURRENT APPLICATION NUMBER: US/09/878,574
;  CURRENT FILING DATE:  2001-12-21
;  PRIOR APPLICATION NUMBER: 09/333,535
;  PRIOR FILING DATE: 1999-06-14
;  NUMBER OF SEQ ID NOS: 15775
; SEQ ID NO 13097
;   LENGTH: 96
;   TYPE: DNA
;   ORGANISM: Glycine max
;   OTHER INFORMATION: Clone ID: 701066628H1
US-09-878-574-13097

  Query Match             66.0%;  Score 13.2;  DB 10;  Length 96;
  Best Local Similarity   83.3%;  Pred. No. 1e+03;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 GATCTTGAACTTCCTCAT 19
           ||||||| || |||| ||
Db       2 GATCTTGCACGTCCTGAT 19


RESULT 13
US-09-231-235-39
; Sequence 39, Application US/09231235
; Patent No. US20020048805A1
; GENERAL INFORMATION:
;  APPLICANT: Johnston, Julie C.
;  APPLICANT: Sauter, Sybille L.
;  APPLICANT: Hsu, David
;  APPLICANT: Sheridan, Philip Lee
;  APPLICANT: Hardy, Steven
;  APPLICANT: Dubensky, Thomas
;  APPLICANT: Yee, Jiing-Kuan
;  TITLE OF INVENTION: FELINE IMMUNODEFICIENCY VIRUS GENE THERAPY VECTORS
;  FILE REFERENCE: 930049.467
;  CURRENT APPLICATION NUMBER: US/09/231,235
;  CURRENT FILING DATE: 1999-01-15
;  NUMBER OF SEQ ID NOS: 63
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 39
;   LENGTH: 21
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Primer
;   FEATURE:
;   OTHER INFORMATION: Primer Used in Example 5B and 5C
US-09-231-235-39

  Query Match             65.0%;  Score 13;  DB 10;  Length 21;
  Best Local Similarity   100.0%;  Pred. No. 1e+03;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       6 TTGAACTTCCTCA 18
           |||||||||||||
Db       1 TTGAACTTCCTCA 13


RESULT 14
US-09-797-518A-39
; Sequence 39, Application US/09797518A
; Patent No. US20020068354A1
; GENERAL INFORMATION:
;  APPLICANT: Johnston, Julie C.
;  APPLICANT:  Sauter, Sybille L.
;  APPLICANT:  Hsu, David
;  APPLICANT:  Sheridan, Philip Lee
;  APPLICANT:  Hardy, Steven
;  APPLICANT:  Dubensky, Thomas
;  APPLICANT:  Yee, Jiing-Kuan
;  TITLE OF INVENTION: FELINE IMMUNODEFICIENCY VIRUS GENE THERAPY VECTORS
;  FILE REFERENCE: 930049.467
;  CURRENT APPLICATION NUMBER: US/09/797,518A
;  CURRENT FILING DATE:  2001-03-01
;  PRIOR APPLICATION NUMBER: 09/231,235
;  PRIOR FILING DATE: 1999-01-15
;  NUMBER OF SEQ ID NOS: 63
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 39
;   LENGTH: 21
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Primer
;   OTHER INFORMATION: Primer Used in Example 5B and 5C
US-09-797-518A-39

  Query Match             65.0%;  Score 13;  DB 10;  Length 21;
  Best Local Similarity   100.0%;  Pred. No. 1e+03;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       6 TTGAACTTCCTCA 18
           |||||||||||||
Db       1 TTGAACTTCCTCA 13


RESULT 15
US-09-231-235-9
; Sequence 9, Application US/09231235
; Patent No. US20020048805A1
; GENERAL INFORMATION:
;  APPLICANT: Johnston, Julie C.
;  APPLICANT: Sauter, Sybille L.
;  APPLICANT: Hsu, David
;  APPLICANT: Sheridan, Philip Lee
;  APPLICANT: Hardy, Steven
;  APPLICANT: Dubensky, Thomas
;  APPLICANT: Yee, Jiing-Kuan
;  TITLE OF INVENTION: FELINE IMMUNODEFICIENCY VIRUS GENE THERAPY VECTORS
;  FILE REFERENCE: 930049.467
;  CURRENT APPLICATION NUMBER: US/09/231,235
;  CURRENT FILING DATE: 1999-01-15
;  NUMBER OF SEQ ID NOS: 63
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
;   LENGTH: 33
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Primer
;   FEATURE:
;   OTHER INFORMATION: Primer Used in Examples 1B and 2B
US-09-231-235-9
```

Query Match 65.0%; Score 13; DB 10; Length 33;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       6 TTGAACTTCCTCA 18
           |||||||||||||
Db      16 TTGAACTTCCTCA 28
```

Search completed: November 23, 2002, 07:10:41
Job time : 16.8 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 26, 2002, 04:16:10 ; Search time 798.5 Seconds
(without alignments)
405.648 Million cell updates/sec

Title: US-09-296-264-25
Perfect score: 20
Sequence: 1 cgatcttgaacttcctcatg 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 357874

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	15.2	76.0	75	13	BI417245	BI417245 LjNEST16e
c	2	15.2	76.0	100	12	BG153998	BG153998 246 L1N01
	3	14.8	74.0	86	9	AA574367	AA574367 nf46b08.s
	4	14.2	71.0	86	9	AI904246	AI904246 PM-BT046-
c	5	13.8	69.0	67	9	AA869223	AA869223 vq49b08.r
c	6	13.8	69.0	76	9	AI120991	AI120991 ub75f06.r
c	7	13.8	69.0	83	10	AW333760	AW333760 S25G3 AGS
	8	13.6	68.0	78	9	AI906723	AI906723 QV-BT124-
c	9	13.6	68.0	92	14	R84789	R84789 yt66e06.r1
c	10	13.6	68.0	97	9	AL817563	AL817563 AL817563
	11	13.4	67.0	40	17	AZ854351	AZ854351 2M0157H21
c	12	13.4	67.0	99	9	AA807995	AA807995 nu98h12.s
c	13	13.2	66.0	26	14	H49860	H49860 yo24e12.s1
c	14	13.2	66.0	41	17	AZ779709	AZ779709 2M0016O20
c	15	13.2	66.0	46	10	AV836228	AV836228 AV836228
	16	13.2	66.0	56	17	BH628418	BH628418 1007079H0
	17	13.2	66.0	64	17	BH814411	BH814411 SALK_0663
	18	13.2	66.0	66	14	C01653	C01653 HUMGS000866
	19	13.2	66.0	82	14	W18704	W18704 mc02c06.r1
c	20	13.2	66.0	93	17	AZ462019	AZ462019 1M0269E07
c	21	13.2	66.0	98	14	BQ822958	BQ822958 1030105B0
	22	12.8	64.0	55	17	AZ612992	AZ612992 1M0441H08
c	23	12.8	64.0	56	10	AW444409	AW444409 AB573 Pri
	24	12.8	64.0	65	10	BE317610	BE317610 NF054C01L
	25	12.8	64.0	71	12	BF541215	BF541215 602068869
c	26	12.8	64.0	74	17	BH855052	BH855052 SALK_0866
c	27	12.8	64.0	80	13	BJ060881	BJ060881 BJ060881
c	28	12.8	64.0	82	17	AZ453529	AZ453529 1M0255A05
c	29	12.8	64.0	86	9	AI973980	AI973980 sd14g01.y
	30	12.8	64.0	90	9	AA668534	AA668534 ac49h02.s
	31	12.8	64.0	93	17	AZ430048	AZ430048 1M0214P17
	32	12.8	64.0	95	9	AL844391	AL844391 AL844391
c	33	12.8	64.0	96	17	AZ438315	AZ438315 1M0228P14
c	34	12.8	64.0	100	9	AA856059	AA856059 vw82c04.r
	35	12.8	64.0	100	10	AW844352	AW844352 RC2-CN005
	36	12.6	63.0	51	14	BQ741582	BQ741582 saq20d02.
	37	12.6	63.0	52	10	BE057696	BE057696 sn05h05.y
	38	12.6	63.0	55	14	C00381	C00381 HUMGS000398
c	39	12.6	63.0	57	14	H26114	H26114 yl48a11.r1
c	40	12.6	63.0	68	9	AA809731	AA809731 nz99e04.s
	41	12.6	63.0	68	13	BM142220	BM142220 if32c03.y
	42	12.6	63.0	69	9	AV200239	AV200239 AV200239
	43	12.6	63.0	70	17	AL751609	AL751609 Arabidops
c	44	12.6	63.0	71	17	BH217655	BH217655 1006057D1
c	45	12.6	63.0	77	17	BH861731	BH861731 SALK_0879

ALIGNMENTS

RESULT 1
BI417245/c
LOCUS BI417245 75 bp mRNA linear EST 15-AUG-2001
DEFINITION LjNEST16e6r Lotus japonicus nodule library 5 and 7 week-old Lotus japonicus cDNA 5', mRNA sequence.
ACCESSION BI417245
VERSION BI417245.1 GI:15188268
KEYWORDS EST.
SOURCE Lotus japonicus.
ORGANISM Lotus japonicus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Loteae; Lotus.
REFERENCE 1 (bases 1 to 75)
AUTHORS Colebatch,G., Freund,S., Trevaskis,B and Udvardi,M.
TITLE Lotus japonicus root nodule ESTs: tools for functional genomics
JOURNAL Unpublished (2000)
COMMENT Contact: Udvardi MK
Molecular Plant Nutrition
Max Planck Institute of Molecular Plant Physiology
Am Muehlenberg 1, 14476 Golm, Germany
Fax: 49 331 567 8250
Email: udvardi@mpimp-golm.mpg.de
Seq primer: T7
High quality sequence stop: 75.
FEATURES Location/Qualifiers
source 1. .75
/organism="Lotus japonicus"

/cultivar="Gifu (B-129)"
/db_xref="taxon:34305"
/clone_lib="Lotus japonicus nodule library 5 and 7 week-old"
/dev_stage="5 and 7 week-old plants"
/note="Organ: Nodule; Vector: pSPORT1; Site_1: SalI; Site_2: NotI; The library was prepared using mRNA extracted from nodules of 5 and 7 week-old Lotus plants. Nodules were induced by, and contained Mesorhizobium strain R7A."
BASE COUNT 19 a 17 c 21 g 18 t
ORIGIN

Query Match 76.0%; Score 15.2; DB 13; Length 75;
Best Local Similarity 85.0%; Pred. No. 6.3e+03;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CGATCTTGAACTTCCTCATG 20
||||||||| ||| | ||||
Db 39 CGATCTTGATCTTGCCCATG 20

RESULT 2
BG153998/c
LOCUS BG153998 100 bp mRNA linear EST 05-FEB-2001
DEFINITION 246 L1N01 Lupinus luteus cDNA, mRNA sequence.
ACCESSION BG153998
VERSION BG153998.1 GI:12666028
KEYWORDS EST.
SOURCE yellow lupine.
ORGANISM Lupinus luteus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Genisteae; Lupinus.
REFERENCE 1 (bases 1 to 100)
AUTHORS Podkowinski,J., Kisiel,A., Grabowska,B. and Legocki,A.B.
TITLE ESTs from the early nodules library of Lupinus luteus
JOURNAL Unpublished (2000)
COMMENT Contact: Podkowinski J
Laboratory of Plant Molecular Biology
Institute of Bioorganic Chemistry of Polish Academy of Sciences
Noskowskiego 12/14, 61-704 Poznan, Poland
Tel: 0048 61 8528919
Fax: 0048 61 8520532
Email: jantsp@ibch.poznan.pl
POLYA=Yes.
FEATURES Location/Qualifiers
source 1. .100
/organism="Lupinus luteus"
/cultivar="Ventus"
/db_xref="taxon:3873"
/clone_lib="L1N01"
/tissue_type="roots with young nodules"
/dev_stage="roots with immature nodules harvested seven days post inoculation with Bradyrhizobium sp. WM9 (Lupinus)"
/lab_host="E. coli strain SOLR"
/note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2: XhoI; cDNA was prepared from polyA+ enriched RNA from roots with developing nodules harvested seven days post inoculation with Bradyrhizobium sp. WM9 (Lupinus). The cDNA was directionally ligated into the Uni-ZAP XR vector from Stratagene and packaged using Gigapack III Gold packaging extracts. Plasmids containing cDNA inserts were excised from the recombinant lambda-ZAP phage using Ex-Assist helper phage and propagated in SOLR cells."
BASE COUNT 28 a 24 c 19 g 29 t
ORIGIN

Query Match 76.0%; Score 15.2; DB 12; Length 100;
Best Local Similarity 85.0%; Pred. No. 6.9e+03;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CGATCTTGAACTTCCTCATG 20
||||||||| ||| | ||||
Db 65 CGATCTTGATCTTGCCCATG 46

RESULT 3
AA574367
LOCUS AA574367 86 bp mRNA linear EST 12-SEP-1997
DEFINITION nf46b08.s1 NCI_CGAP_Pr2 Homo sapiens cDNA clone IMAGE:916791 similar to gb:M90356_cds1 TRANSCRIPTION FACTOR BTF3 (HUMAN);, mRNA sequence.
ACCESSION AA574367
VERSION AA574367.1 GI:2348882
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 86)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: W. Marston Linehan, M.D., Rodrigo Chuaqui, M.D., Michael Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: David B. Krizman, Ph.D.
cDNA Library Arrayed by: Genome Systems Inc., Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html

Trace considered overall poor quality
Insert Length: 2069 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 1.
FEATURES Location/Qualifiers
source 1. .86
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:916791"
/clone_lib="NCI_CGAP_Pr2"
/sex="Male"
/dev_stage="45 years old"
/lab_host="DH10B"
/note="Vector: pAMP10; Site_1: Not1; Site_2: EcoRI; 1st strand cDNA was primed with oligo(dT)17 on 50 ng of DNAse-treated, total cellular RNA obtained from 5,000-10,000 microdissected preneoplastic cells histologically-determined to be prostatic intraepithelial neoplasia 2 (PIN2) cells. Double-stranded cDNA was ligated to EcoRI adaptors, 5 cycles of PCR applied to the cDNA with an adaptor-specific primer, and the resulting PCR product subcloned into pAMP10 by the UDG-cloning method (Life Technologies). Average insert size is 600 bp. NOTE: Not directionally cloned. This library was constructed by David Krizman."
BASE COUNT 22 a 23 c 15 g 26 t
ORIGIN

Query Match 74.0%; Score 14.8; DB 9; Length 86;
Best Local Similarity 88.9%; Pred. No. 1e+04;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 GATCTTGAACTTCCTCAT 19
||||| ||||||| ||||
Db 13 GATCTGGAACTTCATCAT 30

RESULT 4

```
AI904246
LOCUS       AI904246                  86 bp    mRNA    linear   EST 30-MAR-2000
DEFINITION  PM-BT046-120199-200 BT046 Homo sapiens cDNA, mRNA sequence.
ACCESSION   AI904246
VERSION     AI904246.1  GI:6494633
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 86)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
            ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
  MEDLINE   20202663
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/seq/gethtml.pl?tl=PM&t2=PM-BT046-200.html
            &t3=120199&t4=1)
            Seq primer: puc 18 forward.
FEATURES             Location/Qualifiers
     source          1. .86
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="BT046"
                     /sex="female"
                     /dev_stage="Adult"
                     /note="Organ: breast; Vector: puc18; Site_1: SmaI; Site_2:
                     SmaI; A mini-library was made by cloning products derived
                     from  ORESTES PCR (U.S. Letters Patent application No. 196
                     ,716 -  Ludwig Institute for Cancer Research) profiles
                     into  the pUC 18 vector. Reverse transcription of tissue
                     mRNA and cDNA amplification were performed under low
                     stringency conditions."
BASE COUNT       12 a     23 c     33 g     18 t
ORIGIN

  Query Match             71.0%;  Score 14.2;  DB 9;  Length 86;
  Best Local Similarity   84.2%;  Pred. No. 1.9e+04;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 GATCTTGAACTTCCTCATG 20
          |||||||| ||  ||||||
Db     11 GATCTTGATCTGTCTCATG 29


RESULT 5
AA869223/c
LOCUS       AA869223                  67 bp    mRNA    linear   EST 16-MAR-1998
DEFINITION  vq49b08.r1 Barstead bowel MPLRB9 Mus musculus cDNA clone
            IMAGE:1097559 5' similar to SW:UBCA_HUMAN P49459
            UBIQUITIN-CONJUGATING ENZYME E2-17 KD ;, mRNA sequence.
ACCESSION   AA869223
VERSION     AA869223.1  GI:2964668
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 67)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:603791
            Trace considered overall poor quality
            Possible reversed clone: similarity on wrong strand
            Seq primer: -28m13 rev2 ET from Amersham
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .67
                     /organism="Mus musculus"
                     /strain="FVB/N"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:1097559"
                     /clone_lib="Barstead bowel MPLRB9"
                     /tissue_type="bowel"
                     /dev_stage="8 weeks"
                     /lab_host="DH10B"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: EcoRI; Site_2: NotI; 1st strand cDNA
                     was primed with a Not I - oligo(dT) primer [5'
                     TGTTACGAATCTGAAGTGGGAGCGGCCGCCCTTTTTTTTTTTTTTTTTTTTTTT
                     3']; double-stranded cDNA was ligated to Eco RI adaptors
                     [AATTCGGATCCTTG], digested with Not I and cloned into the
                     Not I and Eco RI sites of  the modified pT7T3 vector.
                     Source irradiated bowel harvested 72 hours  after
                     irradiation (1400 Gys).  Library constructed by Bob
                     Barstead.  "
BASE COUNT       14 a     16 c     26 g     11 t
ORIGIN

  Query Match             69.0%;  Score 13.8;  DB 9;  Length 67;
  Best Local Similarity   88.2%;  Pred. No. 2.6e+04;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      4 TCTTGAACTTCCTCATG 20
          ||||||| | |||||||
Db     29 TCTTGAAGTCCCTCATG 13


RESULT 6
AI120991/c
LOCUS       AI120991                  76 bp    mRNA    linear   EST 02-SEP-1998
DEFINITION  ub75f06.r1 Soares_mammary_gland_NMLMG Mus musculus cDNA clone
            IMAGE:1383587 5', mRNA sequence.
ACCESSION   AI120991
VERSION     AI120991.1  GI:3521315
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 76)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
```

```
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:906055
            Seq primer: -28m13 rev2 ET from Amersham
            High quality sequence stop: 69.
FEATURES             Location/Qualifiers
     source          1. .76
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:1383587"
                     /clone_lib="Soares_mammary_gland_NMLMG"
                     /sex="female (lactating)"
                     /tissue_type="mammary gland"
                     /lab_host="DH10B"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; 1st strand cDNA was prepared from mammary
                     gland tissue from a lactating female, and was then primed
                     with a Not I - oligo(dT) primer. Double-stranded  cDNA was
                     ligated to Eco RI adaptors (Pharmacia), digested with Not
                     I  and cloned into the Not I and Eco RI sites of the
                     modified pT7T3 vector.  Library is normalized.  Library
                     was constructed by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT       26 a      6 c     11 g     33 t
ORIGIN

  Query Match             69.0%;  Score 13.8;  DB 9;  Length 76;
  Best Local Similarity   88.2%;  Pred. No. 2.7e+04;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       3 ATCTTGAACTTCCTCAT 19
           ||||| || ||||||||
Db      37 ATCTTAAAATTCCTCAT 21


RESULT 7
AW333760/c
LOCUS       AW333760                 83 bp    mRNA    linear   EST 31-JAN-2000
DEFINITION  S25G3 AGS-1 Pneumocystis carinii f. sp. carinii cDNA 3', mRNA
            sequence.
ACCESSION   AW333760
VERSION     AW333760.1  GI:6830117
KEYWORDS    EST.
SOURCE      Pneumocystis carinii f. sp. carinii.
  ORGANISM  Pneumocystis carinii f. sp. carinii
            Eukaryota; Fungi; Ascomycota; Pneumocystidomycetes;
            Pneumocystidaceae; Pneumocystis.
REFERENCE   1  (bases 1 to 83)
  AUTHORS   Smulian,A.G., Arnold,J., Weise,M., Wunderlich,J., Staben,C., Edman
            ,J.C., Kovacs,J. and Cushion,M.
  TITLE     Expressed sequence tags from Pneumocystis carinii
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Staben C
            School of Biological Sciences
            University of Kentucky
            101 Morgan Building, University of Kentucky, Lexington, KY
            40506-0225, USA
            Tel: 606 257 2161
            Fax: 606 257 1717
            Email: staben@pop.uky.edu.
FEATURES             Location/Qualifiers
     source          1. .83
                     /organism="Pneumocystis carinii f. sp. carinii"
                     /db_xref="taxon:38081"
                     /clone_lib="AGS-1"
                     /lab_host="E. coli"
                     /note="Vector: Lambda ZAP II; Site_1: EcoRI; Site_2: XhoI;
                     P. carinii organisms (3x10e9) from a single rat (99-1-6,
                     sacrificed on 3/17/99)  at Cincinnati VA facilities.
                     Trizol extracted RNA.  Oligo dT priming, standard
                     conditions described by  vendor, Stratagene. Further
                     details see www.uky.edu/Project/Pneumocystis/"
BASE COUNT       27 a     16 c     17 g     23 t
ORIGIN

  Query Match             69.0%;  Score 13.8;  DB 10;  Length 83;
  Best Local Similarity   88.2%;  Pred. No. 2.8e+04;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       4 TCTTGAACTTCCTCATG 20
           |||||||||| |||| |
Db      60 TCTTGAACTTTCTCAGG 44


RESULT 8
AI906723
LOCUS       AI906723                 78 bp    mRNA    linear   EST 30-MAR-2000
DEFINITION  QV-BT124-040399-013 BT124 Homo sapiens cDNA, mRNA sequence.
ACCESSION   AI906723
VERSION     AI906723.1  GI:6497131
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 78)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
            ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
  MEDLINE   20202663
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/seq/gethtml.pl?t1=QV&t2=QV-BT124-013.html
            &t3=040399&t4=1)
            Seq primer: puc 18 forward.
FEATURES             Location/Qualifiers
     source          1. .78
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="BT124"
                     /sex="female"
                     /dev_stage="Adult"
                     /note="Organ: breast; Vector: puc18; Site_1: SmaI; Site_2:
                     SmaI; A mini-library was made by cloning products derived
                     from  ORESTES PCR (U.S. Letters Patent application No. 196
                     ,716 -  Ludwig Institute for Cancer Research) profiles
                     into  the pUC 18 vector. Reverse transcription of tissue
                     mRNA and cDNA amplification were performed under low
                     stringency conditions."
BASE COUNT       17 a     17 c     22 g     22 t
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 9;  Length 78;
  Best Local Similarity   80.0%;  Pred. No. 3.4e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;
```

```
Qy       1 CGATCTTGAACTTCCTCATG 20
           ||||  |||||| || ||||
Db      37 CGATGATGAACTCCCACATG 56


RESULT 9
R84789/c
LOCUS       R84789                   92 bp    mRNA    linear   EST 22-JAN-1999
DEFINITION  yt66e06.r1 Soares retina N2b4HR Homo sapiens cDNA clone
            IMAGE:275411 5' similar to gb:X01677 GLYCERALDEHYDE 3-PHOSPHATE
            DEHYDROGENASE, LIVER (HUMAN);, mRNA sequence.
ACCESSION   R84789
VERSION     R84789.1  GI:943195
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 92)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman
            ,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J.,
            Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston
            ,R., Williamson,A., Wohldmann,P. and Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Insert Size: 1363
            High quality sequence starts: 1
            High quality sequence stops: 1
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Trace considered overall poor quality
            Insert Length: 1363   Std Error: 0.00
            Seq primer: M13RP1
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .92
                     /organism="Homo sapiens"
                     /db_xref="GDB:3849979"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:275411"
                     /clone_lib="Soares retina N2b4HR"
                     /sex="male"
                     /tissue_type="retina"
                     /dev_stage="55 year old"
                     /lab_host="DH10B (ampicillin resistant)"
                     /note="Organ: eye; Vector: pT7T3D (Pharmacia) with a
                     modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCGCTTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was size selected, ligated to Eco RI
                     adapters (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of a modified pT7T3 vector
                     (Pharmacia). The retinas were obtained from a 55 year old
                     Caucasian  and total cellular poly(A)+ RNA was extracted 6
                     hrs after their removal. The retina RNA  was kindly
                     provided by Roderick R. McInnes M.D. Ph.D. from the
                     University  of Toronto. Library constructed by Bento
                     Soares and M.Fatima Bonaldo. "
BASE COUNT       21 a     19 c     28 g     19 t      5 others
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 14;  Length 92;
  Best Local Similarity   80.0%;  Pred. No. 3.6e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 CGATCTTGAACTTCCTCATG 20
           ||| ||| | ||||| ||||
Db      39 CGACCTTCACCTTCCCCATG 20


RESULT 10
AL817563/c
LOCUS       AL817563                 97 bp    mRNA    linear   EST 15-JUL-2002
DEFINITION  AL817563 1:226 Triticum aestivum cDNA clone A05_1226_plate_8, mRNA
            sequence.
ACCESSION   AL817563
VERSION     AL817563.1  GI:21829163
KEYWORDS    EST.
SOURCE      bread wheat.
  ORGANISM  Triticum aestivum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
            ; Triticeae; Triticum.
REFERENCE   1  (bases 1 to 97)
  AUTHORS   Wilson,I., Beswick,R., Shepherd,S., Barker,G., Parker,J., Owen,P.,
            Edwards,D., Coghill,J., Holdsworth,M., Lenton,J., Shewry,P. and
            Edwards,K.
  TITLE     A BBSRC-funded wheat EST resource for the academic community
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Barker G
            Institute of Arable Crop Research
            Long Ashton, Bristol BS41 9AF United Kingdom.
FEATURES             Location/Qualifiers
     source          1. .97
                     /organism="Triticum aestivum"
                     /cultivar="mercia"
                     /db_xref="taxon:4565"
                     /clone="A05_1226_plate_8"
                     /clone_lib="1:226"
                     /tissue_type="embryo"
                     /dev_stage="40 days post anthesis"
BASE COUNT       18 a     39 c     21 g     19 t
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 9;  Length 97;
  Best Local Similarity   80.0%;  Pred. No. 3.7e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 CGATCTTGAACTTCCTCATG 20
           ||||||| | ||| | ||||
Db      84 CGATCTTAATCTTGCCCATG 65


RESULT 11
AZ854351
LOCUS       AZ854351                 40 bp    DNA     linear   GSS 21-FEB-2001
DEFINITION  2M0157H21R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC2M0157H21 R, DNA sequence.
ACCESSION   AZ854351
VERSION     AZ854351.1  GI:13043380
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 40)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
```

```
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0157  row: H  column: 21
            Seq primer: CACACAGGAAACAGCTATGACC
            Class: plasmid ends
            High quality sequence stop: 40.
FEATURES              Location/Qualifiers
     source           1. .40
                      /organism="Mus musculus"
                      /strain="C57BL/6J"
                      /db_xref="taxon:10090"
                      /clone="UUGC2M0157H21"
                      /clone_lib="Mouse 10kb plasmid UUGC1M library"
                      /sex="Male"
                      /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                      /note="Vector: PWD42nv; Purified genomic DNA from M.
                      musculus C57BL/6J (male) was obtained  from the Jackson
                      Laboratory Mouse DNA Resource
                      (http://www.jax.org/resources/documents/dnares/). The DNA
                      was hydrodynamically sheared by repeated passage through a
                      0.005 inch orifice at constant velocity. The sheared DNA
                      was blunt end-repaired with T4 DNA polymerase and T4
                      polynucleotide kinase. Adaptor oligonucleotides were
                      ligated to the blunt ends in high  molar excess. The
                      adaptored DNA was purified and size-selected for a 9.5 to
                      10.5 kb range using preparative agarose gel
                      electrophoresis. Vector DNA was prepared from a derivative
                      of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                      inducible derivative of plasmid R1. The vector was ligated
                      with adaptors complementary to the insert adaptors and
                      purified. The sheared, adaptored mouse DNA was annealed to
                      adaptored vector DNA, and transformed into
                      chemically-competent E. coli XL10-Gold (Stratagene) cells
                      and selected for ampicillin resistance."
BASE COUNT        8 a      7 c     12 g     13 t
ORIGIN

  Query Match             67.0%;  Score 13.4;  DB 17;  Length 40;
  Best Local Similarity   93.3%;  Pred. No. 3.3e+04;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 GATCTTGAACTTCCT 16
           | ||||||||||||||
Db      12 GGTCTTGAACTTCCT 26


RESULT 12
AA807995/c
LOCUS       AA807995                 99 bp    mRNA    linear   EST 12-FEB-1998
DEFINITION  nu98h12.s1 NCI_CGAP_Pr22 Homo sapiens cDNA clone IMAGE:1218791 3'
            similar to gb:V00567 BETA-2-MICROGLOBULIN PRECURSOR (HUMAN);, mRNA
            sequence.
ACCESSION   AA807995
VERSION     AA807995.1  GI:2877401
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 99)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Michael J. Brownstein, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html

            Trace considered overall poor quality
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 1.
FEATURES              Location/Qualifiers
     source           1. .99
                      /organism="Homo sapiens"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:1218791"
                      /clone_lib="NCI_CGAP_Pr22"
                      /sex="male"
                      /tissue_type="normal prostate"
                      /lab_host="DH10B"
                      /note="Organ: prostate; Vector: pT7T3D-Pac (Pharmacia)
                      with a modified polylinker; 1st strand cDNA was prepared
                      from normal prostate bulk tissue, and was then primed with
                      a Not I - oligo(dT) primer. Double-stranded cDNA was
                      ligated to Eco RI adaptors (Pharmacia), digested with Not
                      I and cloned into the Not I  and Eco RI sites of the
                      modified pT7T3 vector.  Library is normalized, and was
                      constructed by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT       30 a     27 c     24 g     18 t
ORIGIN

  Query Match             67.0%;  Score 13.4;  DB 9;  Length 99;
  Best Local Similarity   93.3%;  Pred. No. 4.5e+04;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       5 CTTGAACTTCCTCAT 19
           | ||||||||||||||
Db      75 CCTGAACTTCCTCAT 61


RESULT 13
H49860/c
LOCUS       H49860                   26 bp    mRNA    linear   EST 18-SEP-1995
DEFINITION  yo24e12.s1 Soares adult brain N2b5HB55Y Homo sapiens cDNA clone
            IMAGE:178894 3' similar to gb:M28212 RAS-RELATED PROTEIN RAB-6
            (HUMAN);, mRNA sequence.
ACCESSION   H49860
VERSION     H49860.1  GI:989701
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 26)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman
            ,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J.,
            Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston
            ,R., Williamson,A., Wohldmann,P. and Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Insert Size: 1595
            High quality sequence starts: 1
            High quality sequence stops: 1
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Trace considered overall poor quality
            Insert Length: 1595   Std Error: 0.00
            Seq primer: Promega -21m13
            High quality sequence stop: 1.
FEATURES              Location/Qualifiers
```

```
     source          1. .26
                     /organism="Homo sapiens"
                     /db_xref="GDB:3841090"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:178894"
                     /clone_lib="Soares adult brain N2b5HB55Y"
                     /sex="Male"
                     /dev_stage="55-year old"
                     /lab_host="DH10B (ampicillin resistant)"
                     /note="Organ: brain; Vector: pT7T3D (Pharmacia) with a
                     modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCGCTTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was size selected, ligated to Eco RI
                     adapters (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of a modified pT7T3 vector
                     (Pharmacia). Library went through one round of
                     normalization to a Cot = 53. Library constructed by Bento
                     Soares and M.Fatima Bonaldo. The adult brain RNA was
                     provided by Dr. Donald H. Gilden. Tissue was acquired
                     17-18 hours after death which occurred in consequence of a
                     ruptured aortic aneurysm. RNA was prepared from a pool of
                     tissues representing the following areas of the brain:
                     frontal, parietal, temporal and occipital cortex from the
                     left and right hemispheres, subcortical white matter,
                     basal ganglia, thalamus, cerebellum, midbrain, pons and
                     medulla."
BASE COUNT        8 a      5 c      7 g      6 t
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 14;  Length 26;
  Best Local Similarity   83.3%;  Pred. No. 3.5e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 CGATCTTGAACTTCCTCA 18
           | | |||||| |||||||
Db      25 CCAGCTTGAATTTCCTCA 8


RESULT 14
AZ779709/c
LOCUS       AZ779709                 41 bp    DNA     linear   GSS 16-FEB-2001
DEFINITION  2M0016O20F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC2M0016O20 F, DNA sequence.
ACCESSION   AZ779709
VERSION     AZ779709.1  GI:12910639
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 41)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0016  row: O  column: 20
            Seq primer: CGTTGTAAAACGACGGCCAGT
            Class: plasmid ends
            High quality sequence stop: 41.
FEATURES             Location/Qualifiers
```

```
     source          1. .41
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC2M0016O20"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
BASE COUNT       13 a      2 c     11 g     15 t
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 17;  Length 41;
  Best Local Similarity   83.3%;  Pred. No. 4.1e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       3 ATCTTGAACTTCCTCATG 20
           ||||| |||| ||| |||
Db      32 ATCTTAAACTACCTAATG 15


RESULT 15
AV836228/c
LOCUS       AV836228                 46 bp    mRNA    linear   EST 22-JUN-2001
DEFINITION  AV836228 K. Sato unpublished cDNA library: Hordeum vulgare subsp.
            vulgare seedling leaves second leaf stage Hordeum vulgare subsp.
            vulgare cDNA clone basd2c03, mRNA sequence.
ACCESSION   AV836228
VERSION     AV836228.1  GI:14528317
KEYWORDS    EST.
SOURCE      Hordeum vulgare subsp. vulgare.
  ORGANISM  Hordeum vulgare subsp. vulgare
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
            ; Triticeae; Hordeum.
REFERENCE   1  (bases 1 to 46)
  AUTHORS   Sato,K.
  TITLE     Barley EST sequencing project in NIG and Okayama Univ
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Kazuhiro Sato
            Research Institute for Bioresources
            Okayama University, Barley Germplasm Center
            Chuo 2-20-1, Kurashiki, Okayama 710-0046, Japan
            Email: kazsato@rib.okayama-u.ac.jp,
            URL:http://www.rib.okayama-u.ac.jp/barley/
            Sato,K., Saisho,D., Takeda,K., Shini,T. and Kohara,Y. Direct
            submission;
            database:http://www.shigen.nig.ac.jp/barley/Barley.html.
FEATURES             Location/Qualifiers
     source          1. .46
                     /organism="Hordeum vulgare subsp. vulgare"
                     /cultivar="Haruna Nijo"
                     /db_xref="taxon:112509"
                     /clone="basd2c03"
                     /clone_lib="K. Sato unpublished cDNA library: Hordeum
```

```
                    vulgare subsp. vulgare seedling leaves second leaf stage"
                    /tissue_type="seedling leaves"
                    /dev_stage="second leaf stage"
BASE COUNT       11 a     15 c      9 g     10 t      1 others
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 10;  Length 46;
  Best Local Similarity   78.9%;  Pred. No. 4.3e+04;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 GATCTTGAACTTCCTCATG 20
           | ||| ||||||| | |||
Db      29 GGTCTNGAACTTCTTGATG 11


Search completed: November 26, 2002, 17:57:52
Job time : 809.5 secs
```

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 3, 2002, 18:14:22 ; Search time 351.3 Seconds
(without alignments)
1656.863 Million cell updates/sec

Title: US-09-296-264-26
Perfect score: 20
Sequence: 1 cctgtgagctggaagtcatc 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 995600

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
c	1	14.4	72.0	28	6	AR161733	AR161733 Sequence
c	2	14.2	71.0	51	6	AX157393	AX157393 Sequence
c	3	14.2	71.0	51	6	AX157394	AX157394 Sequence
	4	14.2	71.0	51	6	AX160799	AX160799 Sequence
	5	14.2	71.0	51	6	AX160801	AX160801 Sequence
	6	14.2	71.0	51	6	AX160802	AX160802 Sequence
c	7	13.8	69.0	21	6	I34618	I34618 Sequence 5
	8	13.8	69.0	28	6	BD013011	BD013011 Novel G p
	9	13.8	69.0	28	23	BD010044	Bd010044 Novel G p
	10	13.8	69.0	51	6	AX160797	AX160797 Sequence
	11	13.8	69.0	51	6	AX160798	AX160798 Sequence
c	12	13.6	68.0	37	6	AX357491	AX357491 Sequence
c	13	13.4	67.0	24	6	AX369299	AX369299 Sequence
c	14	13.4	67.0	65	10	AF357505	AF357505 Mus muscu
c	15	13.2	66.0	43	6	A11806	A11806 Trp-6. 12/1
	16	13.2	66.0	90	6	AX241095	AX241095 Sequence
c	17	13	65.0	20	6	AX099782	AX099782 Sequence
c	18	12.8	64.0	20	6	AR169946	AR169946 Sequence
c	19	12.8	64.0	20	6	E16578	E16578 PCR primer
c	20	12.8	64.0	34	6	I29828	I29828 Sequence 14
	21	12.8	64.0	80	6	AX322451	AX322451 Sequence
	22	12.8	64.0	84	11	MMSTS141	Z36680 M.musculus
c	23	12.6	63.0	21	6	AX418162	AX418162 Sequence
	24	12.6	63.0	26	6	AX477108	AX477108 Sequence
c	25	12.6	63.0	29	6	AR138995	AR138995 Sequence
c	26	12.6	63.0	50	6	AX165827	AX165827 Sequence
	27	12.6	63.0	51	6	AX160800	AX160800 Sequence
	28	12.6	63.0	51	6	AX161131	AX161131 Sequence
	29	12.6	63.0	51	6	AX161132	AX161132 Sequence
	30	12.6	63.0	59	6	I01083	I01083 Sequence 1
c	31	12.6	63.0	62	6	A28457	A28457 Linker 6 DN
c	32	12.6	63.0	62	9	AB02076S05	AB020768 Homo sapi
	33	12.6	63.0	66	6	I01084	I01084 Sequence 2
	34	12.6	63.0	70	6	A28270	A28270 linker 2 DN
	35	12.6	63.0	70	6	A28458	A28458 Linker 6 DN
	36	12.6	63.0	72	6	AR129235	AR129235 Sequence
	37	12.6	63.0	83	10	AY041655	AY041655 Oecomys b
	38	12.6	63.0	85	6	AR178146	AR178146 Sequence
	39	12.6	63.0	94	10	AY041899	AY041899 Neacomys
c	40	12.6	63.0	96	9	HUMATPAA2	M16794 Homo sapien
c	41	12.4	62.0	31	6	AX280484	AX280484 Sequence
	42	12.4	62.0	39	6	A37820	A37820 Sequence 3
c	43	12.4	62.0	47	6	AX194759	AX194759 Sequence
c	44	12.4	62.0	51	6	AX165225	AX165225 Sequence
	45	12.4	62.0	66	10	AF193154	AF193154 Mus muscu

ALIGNMENTS

RESULT 1
AR161733/c
LOCUS AR161733 28 bp DNA linear PAT 17-OCT-2001
DEFINITION Sequence 43 from patent US 6258529.
ACCESSION AR161733
VERSION AR161733.1 GI:16228642
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 28)
AUTHORS Berdoz,J. and Kraehenbuhl,J.-P.
TITLE PCR amplification of rearranged genomic variable regions of immunoglobulin genes
JOURNAL Patent: US 6258529-A 43 10-JUL-2001;

```
FEATURES             Location/Qualifiers
     source          1. .28
                     /organism="unknown"
BASE COUNT        6 a      8 c      7 g      7 t
ORIGIN

  Query Match             72.0%;  Score 14.4;  DB 6;  Length 28;
  Best Local Similarity   93.8%;  Pred. No. 1.1e+04;
  Matches   15;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      2 CTGTGAGCTGGAAGTC 17
          ||| ||||||||||||
Db     23 CTGAGAGCTGGAAGTC 8


RESULT 2
AX157393/c
LOCUS       AX157393                 51 bp    DNA     linear   PAT 22-JUN-2001
DEFINITION  Sequence 721 from Patent WO0140521.
ACCESSION   AX157393
VERSION     AX157393.1  GI:14538724
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 51)
  AUTHORS   Shimkets,R.A. and Leach,M.
  TITLE     Nucleic acids containing single nucleotide polymorphisms and
            methods of use thereof
  JOURNAL   Patent: WO 0140521-A 721 07-JUN-2001;
            Curagen Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .51
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     misc_feature    26
                     /note="1 of 2 allelic variants (722 is other entry)
                     Accession number cg21405503"
BASE COUNT       16 a     13 c     12 g     10 t
ORIGIN

  Query Match             71.0%;  Score 14.2;  DB 6;  Length 51;
  Best Local Similarity   84.2%;  Pred. No. 1.4e+04;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 CCTGTGAGCTGGAAGTCAT 19
          ||||||| || |||||| |
Db     21 CCTGTGAACTAGAAGTCTT 3


RESULT 3
AX157394/c
LOCUS       AX157394                 51 bp    DNA     linear   PAT 22-JUN-2001
DEFINITION  Sequence 722 from Patent WO0140521.
ACCESSION   AX157394
VERSION     AX157394.1  GI:14538725
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 51)
  AUTHORS   Shimkets,R.A. and Leach,M.
  TITLE     Nucleic acids containing single nucleotide polymorphisms and
            methods of use thereof
  JOURNAL   Patent: WO 0140521-A 722 07-JUN-2001;
            Curagen Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .51
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     misc_feature    26
                     /note="2 of 2 allelic variants (721 is other entry)
                     Accession number cg21405503"
BASE COUNT       16 a     14 c     11 g     10 t
ORIGIN

  Query Match             71.0%;  Score 14.2;  DB 6;  Length 51;
  Best Local Similarity   84.2%;  Pred. No. 1.4e+04;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 CCTGTGAGCTGGAAGTCAT 19
          ||||||| || |||||| |
Db     21 CCTGTGAACTAGAAGTCTT 3


RESULT 4
AX160799
LOCUS       AX160799                 51 bp    DNA     linear   PAT 22-JUN-2001
DEFINITION  Sequence 4127 from Patent WO0140521.
ACCESSION   AX160799
VERSION     AX160799.1  GI:14542130
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 51)
  AUTHORS   Shimkets,R.A. and Leach,M.
  TITLE     Nucleic acids containing single nucleotide polymorphisms and
            methods of use thereof
  JOURNAL   Patent: WO 0140521-A 4127 07-JUN-2001;
            Curagen Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .51
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     misc_feature    26
                     /note="1 of 2 allelic variants (4128 is other entry)
                     Accession number cg43933539"
BASE COUNT       13 a     14 c     12 g     12 t
ORIGIN

  Query Match             71.0%;  Score 14.2;  DB 6;  Length 51;
  Best Local Similarity   84.2%;  Pred. No. 1.4e+04;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 CTGTGAGCTGGAAGTCATC 20
          ||||||||  ||||||| |
Db     25 CTGTGAGCCAGAAGTCAGC 43


RESULT 5
AX160801
LOCUS       AX160801                 51 bp    DNA     linear   PAT 22-JUN-2001
DEFINITION  Sequence 4129 from Patent WO0140521.
ACCESSION   AX160801
VERSION     AX160801.1  GI:14542132
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 51)
  AUTHORS   Shimkets,R.A. and Leach,M.
  TITLE     Nucleic acids containing single nucleotide polymorphisms and
            methods of use thereof
  JOURNAL   Patent: WO 0140521-A 4129 07-JUN-2001;
            Curagen Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .51
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     misc_feature    26
                     /note="1 of 2 allelic variants (4130 is other entry)
```

```
                     Accession number cg43933539"
BASE COUNT        8 a      14 c      14 g      15 t
ORIGIN

  Query Match             71.0%;  Score 14.2;  DB 6;  Length 51;
  Best Local Similarity   84.2%;  Pred. No. 1.4e+04;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 CTGTGAGCTGGAAGTCATC 20
          ||||||||  ||||||| |
Db      4 CTGTGAGCCAGAAGTCAGC 22


RESULT 6
AX160802
LOCUS       AX160802                  51 bp    DNA     linear   PAT 22-JUN-2001
DEFINITION  Sequence 4130 from Patent WO0140521.
ACCESSION   AX160802
VERSION     AX160802.1  GI:14542133
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 51)
  AUTHORS   Shimkets,R.A. and Leach,M.
  TITLE     Nucleic acids containing single nucleotide polymorphisms and
            methods of use thereof
  JOURNAL   Patent: WO 0140521-A 4130 07-JUN-2001;
            Curagen Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .51
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     misc_feature    26
                     /note="2 of 2 allelic variants (4129 is other entry)
                     Accession number cg43933539"
BASE COUNT        9 a      14 c      13 g      15 t
ORIGIN

  Query Match             71.0%;  Score 14.2;  DB 6;  Length 51;
  Best Local Similarity   84.2%;  Pred. No. 1.4e+04;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 CTGTGAGCTGGAAGTCATC 20
          ||||||||  ||||||| |
Db      4 CTGTGAGCCAGAAGTCAGC 22


RESULT 7
I34618/c
LOCUS       I34618                    21 bp    DNA     linear   PAT 06-FEB-1997
DEFINITION  Sequence 5 from patent US 5596090.
ACCESSION   I34618
VERSION     I34618.1  GI:1825409
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 21)
  AUTHORS   Hoke,G.D., Bradley,M.O., Williams,T.J. and Lee,C.-H.
  TITLE     Antisense oligonucleotides directed against human VCAM-1 RNA
  JOURNAL   Patent: US 5596090-A 5 21-JAN-1997;
FEATURES             Location/Qualifiers
     source          1. .21
                     /organism="unknown"
BASE COUNT        3 a      7 c      3 g      8 t
ORIGIN

  Query Match             69.0%;  Score 13.8;  DB 6;  Length 21;
  Best Local Similarity   88.2%;  Pred. No. 2.4e+04;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      2 CTGTGAGCTGGAAGTCA 18
          ||||||||  |||||||
Db     20 CTGTGAGCAAGAAGTCA 4


RESULT 8
BD013011
LOCUS       BD013011                  28 bp    DNA     linear   PAT 02-AUG-2002
DEFINITION  Novel G protein coupled receptor protein and its DNA.
ACCESSION   BD013011
VERSION     BD013011.1  GI:22093200
KEYWORDS    WO 0116309-A/11.
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 28)
  AUTHORS   Watanabe,T., Terao,Y. and Shintani,Y.
  TITLE     Novel G protein coupled receptor protein and its DNA
  JOURNAL   Patent: WO 0116309-A 11 08-MAR-2001;
            TAKEDA CHEMICAL INDUSTRIES LTD,TAKUYA WATANABE,YASUKO TERAO,
            YASUSHI SHINTANI
COMMENT     OS   Artificial Sequence
            PN   WO 0116309-A/11
            PD   08-MAR-2001
            PF   24-AUG-2000 WO 2000JP005685
            PR   27-AUG-1999 JP 99P   241531,18-JUL-2000 JP 00P   217474 PI
            TAKUYA WATANABE,YASUKO TERAO,YASUSHI SHINTANI PC
            C12N15/09,C07K14/705,C07K16/28,C12N1/21,C12N5/10,C12P21/02, PC
            C12P21/08,
            PC   C12Q1/68,A61K45/00,A61P43/00
            CC
            FH   Key             Location/Qualifiers.
FEATURES             Location/Qualifiers
     source          1. .28
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
BASE COUNT        3 a      5 c      14 g      6 t
ORIGIN

  Query Match             69.0%;  Score 13.8;  DB 6;  Length 28;
  Best Local Similarity   88.2%;  Pred. No. 2.4e+04;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 CCTGTGAGCTGGAAGTC 17
          |||||||||  |||  |||
Db      5 CCTGTGAGCGGGATGTC 21


RESULT 9
BD010044
ID   BD010044   standard; DNA; SYN; 28 BP.
XX
AC   BD010044;
XX
SV   BD010044.1
XX
DT   08-FEB-2002 (Rel. 70, Created)
DT   08-FEB-2002 (Rel. 70, Last updated, Version 1)
XX
DE   Novel G protein coupled receptor protein and its DNA.
XX
KW   JP 03075484-T/11.
XX
OS   synthetic construct
OC   artificial sequence.
XX
RN   [1]
RP   1-28
RA   Watanabe T., Terao Y., Shintani Y.;
RT   "Novel G protein coupled receptor protein and its DNA";
RL   Patent number JP03075484-T/11, 23-FEB-2001.
RL   TAKEDA CHEMICAL INDUSTRIES LTD,TAKUYA WATANABE,YASUKO TERAO, YASUSHI
RL   SHINTANI.
```

```
XX
CC   OS   Artificial Sequence
CC   PN   JP 03075484-T/11
CC   PD   23-FEB-2001
CC   PF   24-AUG-2000 JP 2000005685
CC   PR   27-AUG-1999 JP 99P   241531,18-JUL-2000 JP 00P   217474
CC   PI   TAKUYA WATANABE,YASUKO TERAO,YASUSHI SHINTANI
CC   PC   C12N15/09,C07K14/705,C07K16/28,C12N1/21,C12N5/10,C12P21/02,
CC   PC   C12P21/08,
CC   PC   C12Q1/68,A61K45/00,A61P43/00
CC   CC
CC   FH   Key             Location/Qualifiers
CC   FT   source          1. .28
CC   FT                   /organism="Artificial Sequence"
XX
FH   Key             Location/Qualifiers
FH
FT   source          1. .28
FT                   /db_xref="taxon:32630"
FT                   /organism="synthetic construct"
XX
SQ   Sequence 28 BP; 3 A; 5 C; 14 G; 6 T; 0 other;

  Query Match             69.0%;  Score 13.8;  DB 23;  Length 28;
  Best Local Similarity   88.2%;  Pred. No. 2.4e+04;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 CCTGTGAGCTGGAAGTC 17
          ||||||||| ||| |||
Db      5 CCTGTGAGCGGGATGTC 21


RESULT 10
AX160797
LOCUS       AX160797                 51 bp    DNA     linear   PAT 22-JUN-2001
DEFINITION  Sequence 4125 from Patent WO0140521.
ACCESSION   AX160797
VERSION     AX160797.1  GI:14542128
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 51)
  AUTHORS   Shimkets,R.A. and Leach,M.
  TITLE     Nucleic acids containing single nucleotide polymorphisms and
            methods of use thereof
  JOURNAL   Patent: WO 0140521-A 4125 07-JUN-2001;
            Curagen Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .51
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     misc_feature    26
                     /note="1 of 2 allelic variants (4126 is other entry)
                     Accession number cg43933539"
BASE COUNT       14 a     11 c     12 g     14 t
ORIGIN

  Query Match             69.0%;  Score 13.8;  DB 6;  Length 51;
  Best Local Similarity   88.2%;  Pred. No. 2.4e+04;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      2 CTGTGAGCTGGAAGTCA 18
          |||||||| |||||||
Db     34 CTGTGAGCCAGAAGTCA 50


RESULT 11
AX160798
LOCUS       AX160798                 51 bp    DNA     linear   PAT 22-JUN-2001
DEFINITION  Sequence 4126 from Patent WO0140521.
ACCESSION   AX160798
VERSION     AX160798.1  GI:14542129
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 51)
  AUTHORS   Shimkets,R.A. and Leach,M.
  TITLE     Nucleic acids containing single nucleotide polymorphisms and
            methods of use thereof
  JOURNAL   Patent: WO 0140521-A 4126 07-JUN-2001;
            Curagen Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .51
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     misc_feature    26
                     /note="2 of 2 allelic variants (4125 is other entry)
                     Accession number cg43933539"
BASE COUNT       13 a     11 c     13 g     14 t
ORIGIN

  Query Match             69.0%;  Score 13.8;  DB 6;  Length 51;
  Best Local Similarity   88.2%;  Pred. No. 2.4e+04;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      2 CTGTGAGCTGGAAGTCA 18
          |||||||| |||||||
Db     34 CTGTGAGCCAGAAGTCA 50


RESULT 12
AX357491/c
LOCUS       AX357491                 37 bp    DNA     linear   PAT 13-FEB-2002
DEFINITION  Sequence 4 from Patent WO0190400.
ACCESSION   AX357491
VERSION     AX357491.1  GI:18674547
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Gu,K.
  TITLE     Methods for monitoring production of gene products and uses thereof
  JOURNAL   Patent: WO 0190400-A 4 29-NOV-2001;
            Gu, Kerong (US)
FEATURES             Location/Qualifiers
     source          1. .37
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="Oligonucleotide 3p-kan-DsRed"
BASE COUNT       11 a      6 c     13 g      7 t
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 6;  Length 37;
  Best Local Similarity   80.0%;  Pred. No. 3.1e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 CCTGTGAGCTGGAAGTCATC 20
          |||  ||||| ||| |||||
Db     23 CCTTCGAGCTCGAATTCATC 4


RESULT 13
AX369299/c
LOCUS       AX369299                 24 bp    DNA     linear   PAT 16-FEB-2002
DEFINITION  Sequence 7 from Patent WO0206321.
ACCESSION   AX369299
VERSION     AX369299.1  GI:18857244
KEYWORDS    .
SOURCE      Zea mays.
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
```

```
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1
  AUTHORS   Kaeppler,S.M., Springer,N.M. and Phillips,R.L.
  TITLE     Polycomb gene from maize- zmfie2
  JOURNAL   Patent: WO 0206321-A 7 24-JAN-2002;
            WISCONSIN ALUMNI RESEARCH FOUNDATION (US) ; REGENTS OF THE
            UNIVERSITY OF MINNESOTA (US)
FEATURES             Location/Qualifiers
     source          1. .24
                     /organism="Zea mays"
                     /db_xref="taxon:4577"
BASE COUNT        4 a      9 c      4 g      7 t
ORIGIN

  Query Match           67.0%;  Score 13.4;  DB 6;  Length 24;
  Best Local Similarity   93.3%;  Pred. No. 3.9e+04;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       6 GAGCTGGAAGTCATC 20
           ||||||| ||||||
Db      24 GAGCTGGAGGTCATC 10


RESULT 14
AF357505/c
LOCUS       AF357505                  65 bp    RNA     linear   ROD 06-JUN-2001
DEFINITION  Mus musculus clone MBII-97 miscellaneous RNA, partial sequence.
ACCESSION   AF357505
VERSION     AF357505.1  GI:14277100
KEYWORDS    .
SOURCE      Mus musculus.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 65)
  AUTHORS   Huttenhofer,A., Kiefmann,M., Meier-Ewert,S., O'Brien,J.,
            Lehrach,H., Bachellerie,J.P. and Brosius,J.
  TITLE     RNomics: an experimental approach that identifies 201 candidates
            for novel, small, non-messenger RNAs in mouse
  JOURNAL   EMBO J. 20 (11), 2943-2953 (2001)
  MEDLINE   21280816
   PUBMED   11387227
REFERENCE   2  (bases 1 to 65)
  AUTHORS   Huttenhofer,A., Kiefmann,M., Meier-Ewert,S., O'Brien,J.,
            Lehrach,H., Bachellerie,J.-P. and Brosius,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-MAR-2001) Institute of Experimental Pathology /
            Molecular Neurobiology, ZMBE, University of Muenster,
            Von-Esmarch-Str. 56, Muenster D-48149, Germany
FEATURES             Location/Qualifiers
     source          1. .65
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /clone="MBII-97"
     misc_RNA        <1. .65
                     /product="derived from hnRNA or mRNA fragment, or novel
                     small non-messenger RNA without known sequence-or
                     structural motifs"
                     /note="missing at least 5-10 bases"
BASE COUNT       13 a     15 c     16 g     21 t
ORIGIN

  Query Match           67.0%;  Score 13.4;  DB 10;  Length 65;
  Best Local Similarity   93.3%;  Pred. No. 3.9e+04;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       6 GAGCTGGAAGTCATC 20
           |||||||||| |||
Db      31 GAGCTGGAAGTGATC 17


RESULT 15
A11806/c
LOCUS       A11806                    43 bp    DNA     linear   PAT 01-DEC-1993
DEFINITION  Trp-6.
ACCESSION   A11806
VERSION     A11806.1  GI:492525
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 43)
  AUTHORS   Hauptmann,R., Himmler,A. and Swetly,P.
  TITLE     Horse gamma interferon
  JOURNAL   Patent: EP 0271824-A 31 22-JUN-1988;
            BOEHRINGER INGELHEIM INTERNATIONAL GmbH
FEATURES             Location/Qualifiers
     source          1. .43
                     /organism="unidentified"
                     /db_xref="taxon:32644"
BASE COUNT       11 a     13 c      7 g     12 t
ORIGIN

  Query Match           66.0%;  Score 13.2;  DB 6;  Length 43;
  Best Local Similarity   83.3%;  Pred. No. 5e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       3 TGTGAGCTGGAAGTCATC 20
           | |||||| ||| |||||
Db      19 TATGAGCTCGAATTCATC 2


Search completed: December  3, 2002, 22:23:33
Job time : 357.3 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 06:29:45 ; Search time 94.1 Seconds
(without alignments)
478.639 Million cell updates/sec

Title: US-09-296-264-26
Perfect score: 20
Sequence: 1 cctgtgagctggaagtcatc 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 2390332

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	20	100.0	20	21	AAZ31456	Human neuropilin m
c	2	14.8	74.0	60	24	ABN34924	Human spliced tran
c	3	14.8	74.0	60	24	ABN43595	Human spliced tran
c	4	14.4	72.0	28	17	AAT30493	Primer 5'kappa30 f
c	5	14.4	72.0	77	22	ABA75759	Human foetal liver
c	6	14.4	72.0	77	22	ABA40331	Probe #18797 for g
c	7	14.2	71.0	33	24	AAL44794	Human RNA uncoilin
c	8	14.2	71.0	51	22	AAI73780	Human silent SNP c
c	9	14.2	71.0	51	22	AAI73781	Human silent SNP c
	10	14.2	71.0	51	22	AAI77186	Human silent SNP c
	11	14.2	71.0	51	22	AAI77188	Human silent SNP c
	12	14.2	71.0	51	22	AAI77189	Human silent SNP c
c	13	14.2	71.0	60	24	ABN35100	Human spliced tran
c	14	14.2	71.0	60	24	ABN48570	Human spliced tran
c	15	14.2	71.0	65	24	ABN52112	Mouse spliced tran
	16	13.8	69.0	18	20	AAX25981	Primer Seq2 scAb.
c	17	13.8	69.0	21	18	AAT64323	Antisense oligonuc
c	18	13.8	69.0	21	19	AAV38641	Human ICAM-1, E-se
	19	13.8	69.0	28	22	AAF79511	Human G protein-co
	20	13.8	69.0	28	24	ABL49625	Human G protein-co
	21	13.8	69.0	44	24	ABL53470	PCR primer for con
c	22	13.8	69.0	50	21	AAA96320	Probe for cDNA enc
	23	13.8	69.0	51	22	AAI77184	Human silent SNP c
	24	13.8	69.0	51	22	AAI77185	Human silent SNP c
	25	13.8	69.0	60	24	ABN45637	Human spliced tran
	26	13.8	69.0	65	24	ABN56055	Mouse spliced tran
c	27	13.8	69.0	65	24	ABN56886	Mouse spliced tran
c	28	13.6	68.0	37	24	AAD28541	Yeast DsRed strain
	29	13.6	68.0	60	24	ABN37742	Human spliced tran
c	30	13.6	68.0	60	24	ABN44525	Human spliced tran
c	31	13.4	67.0	20	24	AAL38277	Mouse BH3 interact
c	32	13.4	67.0	24	24	ABK10008	Maize polycomb gen
	33	13.4	67.0	51	22	AAL32714	Human SNP oligonuc
	34	13.4	67.0	60	24	ABN43432	Human spliced tran
	35	13.4	67.0	60	24	ABN47221	Human spliced tran
	36	13.2	66.0	22	24	ABK46916	COX-2 antisense ol
	37	13.2	66.0	27	24	AAL40541	Human ABCB1 gene r
	38	13.2	66.0	27	24	ABK82174	Human ATP-binding
c	39	13.2	66.0	41	24	AAL44796	Human RNA uncoilin
c	40	13.2	66.0	41	24	AAL44797	Human RNA uncoilin
	41	13.2	66.0	60	24	ABN36668	Human spliced tran
c	42	13.2	66.0	73	18	AAV75900	Staphylococcus aur
	43	13.2	66.0	90	22	AAS23652	Tetracycline promo
c	44	13	65.0	20	22	AAD03301	Forward primer #3
	45	13	65.0	60	24	ABN35071	Human spliced tran

ALIGNMENTS

RESULT 1
AAZ31456
ID AAZ31456 standard; DNA; 20 BP.
XX
AC AAZ31456;
XX
DT 07-FEB-2000 (first entry)
XX
DE Human neuropilin mRNA specific antisense oligo GTI3627.
XX
KW Neuropilin; human; growth; metastasis; tumor; neovascularisation;
KW cancer; papilloma; diabetic retinopathy; antisense; ss.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9955855-A2.
XX
PD 04-NOV-1999.
XX
PF 23-APR-1999; 99WO-CA00324.
XX
PR 23-APR-1998; 98US-0082791.
XX
PA (GENE-) GENESENSE TECHNOLOGIES INC.
XX
PI Wright JA, Young AH, Lee YS;
XX
DR WPI; 2000-023357/02.
XX
PT Antisense oligonucleotides that inhibit neuropilin expression, useful
PT for treating cancer -

XX
PS Claim 4; Page 17; 57pp; English.
XX
CC Sequences AAZ31431-460 represent antisense oligonucleotides which
CC inhibit human neuropilin expression. The antisense oligonucleotides can
CC be used to inhibit the growth or metastasis of a mammalian tumor and
CC inhibit neovascularisation. The oligonucleotides may be used to treat
CC various forms of cancers or tumors, such as sarcomas, melanomas,
CC adenomas, carcinomas of solid tissue, hypoxic tumors, squamous cell
CC carcinomas of the mouth, throat, larynx and lung, genitourinary cancers
CC such as cervical and bladder cancer, hematopoietic cancers, colon cancer,
CC breast cancer, pancreatic cancer, renal cancer, brain cancer, skin
CC cancer, liver cancer, head and neck cancers, and nervous system cancers,
CC as well as benign lesions such as papillomas. The methods may be used to
CC treat neovascularisation disorders such as diabetic retinopathy, and
CC retinopathy of prematurity and age related macular degeneration.
XX
SQ Sequence 20 BP; 4 A; 5 C; 6 G; 5 T; 0 other;

Query Match 100.0%; Score 20; DB 21; Length 20;
Best Local Similarity 100.0%; Pred. No. 2.3;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCTGTGAGCTGGAAGTCATC 20
||||||||||||||||||||
Db 1 CCTGTGAGCTGGAAGTCATC 20

RESULT 2
ABN34924/c
ID ABN34924 standard; DNA; 60 BP.
XX
AC ABN34924;
XX
DT 15-JUL-2002 (first entry)
XX
DE Human spliced transcript detection oligonucleotide SEQ ID NO:7672.
XX
KW Human; mouse; rat; splice transcript; detection; RNA transcript;
KW splice variant; transcriptome; oligonucleotide library; ss.
XX
OS Homo sapiens.
XX
PN WO200210449-A2.
XX
PD 07-FEB-2002.
XX
PF 20-JUL-2001; 2001WO-IB01903.
XX
PR 28-JUL-2000; 2000US-221607P.
PR 02-MAY-2001; 2001US-287724P.
XX
PA (COMP-) COMPUGEN INC.
XX
PI Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
XX
DR WPI; 2002-257383/30.
XX
PT New oligonucleotide libraries comprising oligonucleotides which
PT selectively hybridize to mRNAs transcribed from a transcription unit of
PT a genome, useful for detecting tissue-, pathology-, and
PT developmental-specific genes -
XX
PS Example 1; SEQ ID 7672; 47pp; English.
XX
CC The present invention describes oligonucleotide libraries for detecting
CC messenger RNAs that populate a (sub-)transcriptome, where the
CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC transcription units that populate a genome. The library comprises
CC several oligonucleotides, each capable of hybridising selectively to a
CC set of messenger RNAs transcribed from a given transcription unit of
CC the genome, which encodes one or more messenger RNA splice variants.
CC The oligonucleotide libraries are useful for detecting mRNAs from a
CC biological sample, in expression profiling studies, in qualitatively or
CC quantitatively characterising the corresponding transcriptome, and in
CC detecting RNA transcripts and splice variants of human or animal
CC transcriptomes. The libraries may also be used as specialised mini
CC libraries to detect transcripts of a sub-transcriptome under a
CC particular biological or pathological state, and so allowing the
CC detection of tissue- and pathology-specific genes such as those genes
CC only expressed in specific tissue under a specific pathological
CC condition; to detect developmental specific genes; and to detect RNA
CC transcripts and splice variants of a transcriptome of a patient suffering
CC from a particular disorder. ABN27253 to ABN59589 represent
CC oligonucleotide sequences from rats, humans and mice, which are used in
CC the exemplification of the present invention.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 60 BP; 20 A; 16 C; 14 G; 10 T; 0 other;

Query Match 74.0%; Score 14.8; DB 24; Length 60;
Best Local Similarity 88.9%; Pred. No. 8.9e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 CTGTGAGCTGGAAGTCAT 19
|||||||| ||||||| |
Db 28 CTGTGAGCAGGAAGTCCT 11

RESULT 3
ABN43595/c
ID ABN43595 standard; DNA; 60 BP.
XX
AC ABN43595;
XX
DT 15-JUL-2002 (first entry)
XX
DE Human spliced transcript detection oligonucleotide SEQ ID NO:16343.
XX
KW Human; mouse; rat; splice transcript; detection; RNA transcript;
KW splice variant; transcriptome; oligonucleotide library; ss.
XX
OS Homo sapiens.
XX
PN WO200210449-A2.
XX
PD 07-FEB-2002.
XX
PF 20-JUL-2001; 2001WO-IB01903.
XX
PR 28-JUL-2000; 2000US-221607P.
PR 02-MAY-2001; 2001US-287724P.
XX
PA (COMP-) COMPUGEN INC.
XX
PI Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
XX
DR WPI; 2002-257383/30.
XX
PT New oligonucleotide libraries comprising oligonucleotides which
PT selectively hybridize to mRNAs transcribed from a transcription unit of
PT a genome, useful for detecting tissue-, pathology-, and
PT developmental-specific genes -
XX
PS Example 1; SEQ ID 16343; 47pp; English.
XX
CC The present invention describes oligonucleotide libraries for detecting
CC messenger RNAs that populate a (sub-)transcriptome, where the
CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC transcription units that populate a genome. The library comprises
CC several oligonucleotides, each capable of hybridising selectively to a
CC set of messenger RNAs transcribed from a given transcription unit of
CC the genome, which encodes one or more messenger RNA splice variants.
CC The oligonucleotide libraries are useful for detecting mRNAs from a

CC biological sample, in expression profiling studies, in qualitatively or
CC quantitatively characterising the corresponding transcriptome, and in
CC detecting RNA transcripts and splice variants of human or animal
CC transcriptomes. The libraries may also be used as specialised mini
CC libraries to detect transcripts of a sub-transcriptome under a
CC particular biological or pathological state, and so allowing the
CC detection of tissue- and pathology-specific genes such as those genes
CC only expressed in specific tissue under a specific pathological
CC condition; to detect developmental specific genes; and to detect RNA
CC transcripts and splice variants of a transcriptome of a patient suffering
CC from a particular disorder. ABN27253 to ABN59589 represent
CC oligonucleotide sequences from rats, humans and mice, which are used in
CC the exemplification of the present invention.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 60 BP; 15 A; 17 C; 13 G; 15 T; 0 other;

Query Match 74.0%; Score 14.8; DB 24; Length 60;
Best Local Similarity 88.9%; Pred. No. 8.9e+02;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CCTGTGAGCTGGAAGTCA 18
||||||| |||||||| ||
Db 43 CCTGTGTGCTGGAAGCCA 26

RESULT 4
AAT30493/c
ID AAT30493 standard; DNA; 28 BP.
XX
AC AAT30493;
XX
DT 11-FEB-1997 (first entry)
XX
DE Primer 5'kappa30 for HNK-20 Vkappa chain coding sequence.
XX
KW Antibody; HNK-20; variable kappa chain; hybridoma; murine; IgA; mouse;
KW F glycoprotein; respiratory syncytial virus; RSV; constant region gene;
KW chimeric antibody; isotype-switched antibody; therapy; infection; human;
KW pneumonia; bronchiolitis; animal; polymerase chain reaction; primer; PCR;
KW amplify; ss.
XX
OS Synthetic.
XX
PN WO9616974-A1.
XX
PD 06-JUN-1996.
XX
PF 01-DEC-1995; 95WO-US15716.
XX
PR 01-DEC-1994; 94US-0348548.
XX
PA (ORAV-) ORAVAX INC.
XX
PI Berdoz J, Kraehenbuhl J;
XX
DR WPI; 1996-286826/29.
XX
PT DNA encoding variable region of antibody HNK-20 - for treating
PT respiratory syncytial virus infection
XX
PS Example; Page 40; 75pp; English.
XX
CC AAT30459-T30545 represent amplification primers for the coding sequences
CC for the variable chains of an antibody produced by the hybridoma cell
CC line HNK-20. AAT30464-T30498 represent amplification primers for the
CC coding sequence for the variable kappa chain of the HNK-20 antibody.
CC HNK-20 is a murine hybridoma cell line, that produces IgA specific for
CC the F glycoprotein of respiratory syncytial virus (RSV). The variable
CC chain coding sequences (see AAT30456-T30458) were isolated using primers
CC specific for the 5' untranslated region of the variable region, and for
CC the intron downstream of the rearranged J region. The amplified
CC sequences can be inserted into vectors containing heterologous (such as
CC human) constant region genes, for the production of chimeric and
CC isotype-switched antibodies. The antibodies are useful in the treatment
CC and diagnosis of infection by RSV, such as pneumonia and bronchiolitis,
CC in humans and animals. By using genomic DNA as a template, variable
CC region genes can be isolated without producing fragments that have to be
CC adapted for recombinant antibody expression. Also, by using the genomic
CC DNA, no knowledge of the DNA sequence encoding the target variable
CC region is required. Chimeric antibodies produced from the encoded
CC proteins, that contain the constant region of the host being treated, are
CC less likely to cause adverse immune reactions.
XX
SQ Sequence 28 BP; 6 A; 8 C; 7 G; 7 T; 0 other;

Query Match 72.0%; Score 14.4; DB 17; Length 28;
Best Local Similarity 93.8%; Pred. No. 1.3e+03;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 CTGTGAGCTGGAAGTC 17
||| ||||||||||||
Db 23 CTGAGAGCTGGAAGTC 8

RESULT 5
ABA75759/c
ID ABA75759 standard; DNA; 77 BP.
XX
AC ABA75759;
XX
DT 01-FEB-2002 (first entry)
XX
DE Human foetal liver single exon nucleic acid probe #24064.
XX
KW Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
XX
OS Homo sapiens.
XX
PN WO200157277-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00669.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-483447/52.
XX
PT Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human fetal liver -
XX
PS Claim 4; SEQ ID NO 24064; 639pp + sequence listing; English.
XX
CC The invention relates to a single exon nucleic acid probe for
CC measuring human gene expression in a sample derived from human foetal
CC liver. The single exon nucleic acid probes may be used for predicting,
CC measuring and displaying gene expression in samples derived from human
CC fetal liver. The present sequence is a single exon nucleic acid
CC probe of the invention.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX

```
SQ   Sequence 77 BP; 28 A; 14 C; 20 G; 15 T; 0 other;

  Query Match            72.0%;  Score 14.4;  DB 22;  Length 77;
  Best Local Similarity  93.8%;  Pred. No. 1.4e+03;
  Matches   15;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      2 CTGTGAGCTGGAAGTC 17
          |||||| |||||||||
Db     77 CTGTGACCTGGAAGTC 62


RESULT 6
ABA40331/c
ID   ABA40331 standard; DNA; 77 BP.
XX
AC   ABA40331;
XX
DT   23-JAN-2002  (first entry)
XX
DE   Probe #18797 for gene expression analysis in human heart cell sample.
XX
KW   Human; gene expression; heart; microarray; vascular system; probe;
KW   cardiovascular disease; hypertension; cardiac arrhythmia;
KW   congenital heart disease; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157274-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00666.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488899/53.
XX
PT   Single exon nucleic acid probes for analyzing gene expression in human
PT   hearts -
XX
PS   Claim 4; SEQ ID No 18797; 530pp; English.
XX
CC   The present invention relates to single exon nucleic acid probes for
CC   measuring human gene expression in a sample derived from human heart. The
CC   present sequence is one such probe. The probes may be used for
CC   predicting, measuring and displaying gene expression in samples derived
CC   from the human heart via microarrays. By measuring gene expression, the
CC   probes are useful for predicting, diagnosing, grading, staging,
CC   monitoring and prognosing diseases of the human heart and vascular system
CC   e.g. cardiovascular disease, hypertension, cardiac arrhythmias and
CC   congenital heart disease.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 77 BP; 28 A; 14 C; 20 G; 15 T; 0 other;

  Query Match            72.0%;  Score 14.4;  DB 22;  Length 77;
  Best Local Similarity  93.8%;  Pred. No. 1.4e+03;
  Matches   15;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      2 CTGTGAGCTGGAAGTC 17
          |||||| |||||||||
Db     77 CTGTGACCTGGAAGTC 62


RESULT 7
AAL44794/c
ID   AAL44794 standard; DNA; 33 BP.
XX
AC   AAL44794;
XX
DT   16-MAY-2002  (first entry)
XX
DE   Human RNA uncoiling enzyme 2 coding sequence PCR primer #3.
XX
KW   Human; RNA uncoiling enzyme 2; cancer; haemopathy; HIV infection;
KW   immune disease; inflammation; gene therapy; PCR primer; enzyme; ss.
XX
OS   Homo sapiens.
XX
PN   CN1323896-A.
XX
PD   28-NOV-2001.
XX
PF   16-MAY-2000; 2000CN-0115692.
XX
PR   16-MAY-2000; 2000CN-0115692.
XX
PA   (SHAN-) SHANGHAI BODE GENE DEV CO LTD.
XX
PI   Mao Y,  Xie Y;
XX
DR   WPI; 2002-148842/20.
XX
PT   New polypeptide-human RNA uncoiling enzyme 12 and polynucleotides for
PT   coding same -
XX
PS   Example 4; Page 18(Disclosure); 34pp; Chinese.
XX
CC   The present invention provides the protein and coding sequences of human
CC   RNA uncoiling enzyme 2. The sequences can be used in the treatment of
CC   cancer, haemopathy, HIV infection, immune diseases and inflammation. The
CC   present sequence is a PCR primer for the coding sequence of the
CC   invention.
XX
SQ   Sequence 33 BP; 9 A; 11 C; 6 G; 7 T; 0 other;

  Query Match            71.0%;  Score 14.2;  DB 24;  Length 33;
  Best Local Similarity  84.2%;  Pred. No. 1.7e+03;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 CCTGTGAGCTGGAAGTCAT 19
          |||| | ||||| ||||||
Db     28 CCTGGGTGCTGGGAGTCAT 10


RESULT 8
AAI73780/c
ID   AAI73780 standard; DNA; 51 BP.
XX
AC   AAI73780;
XX
DT   09-NOV-2001  (first entry)
XX
DE   Human silent SNP containing nucleic acid SEQ:721.
XX
KW   Human; single nucleotide polymorphism; SNP; genome; gene therapy;
KW   protein therapy; vaccine; probe; diagnostic assay; detection;
KW   quantitation; restorative therapy; polymorphic; ds.
XX
OS   Homo sapiens.
XX
PN   WO200140521-A2.
XX
PD   07-JUN-2001.
```

XX
PF 30-NOV-2000; 2000WO-US32758.
XX
PR 30-NOV-1999; 99US-0168138.
PR 29-NOV-2000; 2000US-0726173.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shimkets RA, Leach M;
XX
DR WPI; 2001-356160/37.
XX
PT Polymorphic nucleic acid sequences, useful in genetic testing and
PT therapy -
XX
PS Claim 1; Page 274; 2653pp; English.
XX
CC AAI73060 to AAI79867 represent isolated human polymorphic polynucleotide
CC sequences (I), which contain single nucleotide polymorphisms (SNPs).
CC AAM53114 to AAM53329 represent peptides related to human polymorphic
CC polynucleotide sequences. The sequences can be used in gene and protein
CC therapy, and in vaccine production. (I) and the polypeptides encoded by
CC them may be used in the prevention, diagnosis and treatment of diseases
CC associated with inappropriate expression of polymorphic polypeptides.
CC For example, (I) may be used to treat disorders by rectifying mutations
CC or deletions in a patient's genome that affect the activity of
CC polypeptides by expressing inactive proteins or to supplement the
CC patients own production of polypeptide. Additionally, (I) and its
CC complementary sequences may also be used as DNA probes in diagnostic
CC assays to detect and quantitate the presence of similar nucleic acids
CC in samples, and therefore which patients may be in need of restorative
CC therapy. The polypeptides encoded by (I) may be used as antigens in the
CC production of antibodies specific for polymorphic polypeptides. The
CC antibodies may also be used to down regulate expression and activity.
CC The antibodies may also be used as diagnostic agents for detecting the
CC presence of polymorphic polypeptides in samples.
XX
SQ Sequence 51 BP; 16 A; 13 C; 12 G; 10 T; 0 other;

Query Match 71.0%; Score 14.2; DB 22; Length 51;
Best Local Similarity 84.2%; Pred. No. 1.7e+03;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CCTGTGAGCTGGAAGTCAT 19
||||||| || |||||| |
Db 21 CCTGTGAACTAGAAGTCTT 3

RESULT 9
AAI73781/c
ID AAI73781 standard; DNA; 51 BP.
XX
AC AAI73781;
XX
DT 09-NOV-2001 (first entry)
XX
DE Human silent SNP containing nucleic acid SEQ:722.
XX
KW Human; single nucleotide polymorphism; SNP; genome; gene therapy;
KW protein therapy; vaccine; probe; diagnostic assay; detection;
KW quantitation; restorative therapy; polymorphic; ds.
XX
OS Homo sapiens.
XX
PN WO200140521-A2.
XX
PD 07-JUN-2001.
XX
PF 30-NOV-2000; 2000WO-US32758.
XX
PR 30-NOV-1999; 99US-0168138.
PR 29-NOV-2000; 2000US-0726173.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shimkets RA, Leach M;
XX
DR WPI; 2001-356160/37.
XX
PT Polymorphic nucleic acid sequences, useful in genetic testing and
PT therapy -
XX
PS Claim 1; Page 275; 2653pp; English.
XX
CC AAI73060 to AAI79867 represent isolated human polymorphic polynucleotide
CC sequences (I), which contain single nucleotide polymorphisms (SNPs).
CC AAM53114 to AAM53329 represent peptides related to human polymorphic
CC polynucleotide sequences. The sequences can be used in gene and protein
CC therapy, and in vaccine production. (I) and the polypeptides encoded by
CC them may be used in the prevention, diagnosis and treatment of diseases
CC associated with inappropriate expression of polymorphic polypeptides.
CC For example, (I) may be used to treat disorders by rectifying mutations
CC or deletions in a patient's genome that affect the activity of
CC polypeptides by expressing inactive proteins or to supplement the
CC patients own production of polypeptide. Additionally, (I) and its
CC complementary sequences may also be used as DNA probes in diagnostic
CC assays to detect and quantitate the presence of similar nucleic acids
CC in samples, and therefore which patients may be in need of restorative
CC therapy. The polypeptides encoded by (I) may be used as antigens in the
CC production of antibodies specific for polymorphic polypeptides. The
CC antibodies may also be used to down regulate expression and activity.
CC The antibodies may also be used as diagnostic agents for detecting the
CC presence of polymorphic polypeptides in samples.
XX
SQ Sequence 51 BP; 16 A; 14 C; 11 G; 10 T; 0 other;

Query Match 71.0%; Score 14.2; DB 22; Length 51;
Best Local Similarity 84.2%; Pred. No. 1.7e+03;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CCTGTGAGCTGGAAGTCAT 19
||||||| || |||||| |
Db 21 CCTGTGAACTAGAAGTCTT 3

RESULT 10
AAI77186
ID AAI77186 standard; DNA; 51 BP.
XX
AC AAI77186;
XX
DT 09-NOV-2001 (first entry)
XX
DE Human silent SNP containing nucleic acid SEQ:4127.
XX
KW Human; single nucleotide polymorphism; SNP; genome; gene therapy;
KW protein therapy; vaccine; probe; diagnostic assay; detection;
KW quantitation; restorative therapy; polymorphic; ds.
XX
OS Homo sapiens.
XX
PN WO200140521-A2.
XX
PD 07-JUN-2001.
XX
PF 30-NOV-2000; 2000WO-US32758.
XX
PR 30-NOV-1999; 99US-0168138.
PR 29-NOV-2000; 2000US-0726173.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shimkets RA, Leach M;
XX
DR WPI; 2001-356160/37.
XX

PT Polymorphic nucleic acid sequences, useful in genetic testing and
PT therapy -
XX
PS Claim 1; Page 1773; 2653pp; English.
XX
CC AAI73060 to AAI79867 represent isolated human polymorphic polynucleotide
CC sequences (I), which contain single nucleotide polymorphisms (SNPs).
CC AAM53114 to AAM53329 represent peptides related to human polymorphic
CC polynucleotide sequences. The sequences can be used in gene and protein
CC therapy, and in vaccine production. (I) and the polypeptides encoded by
CC them may be used in the prevention, diagnosis and treatment of diseases
CC associated with inappropriate expression of polymorphic polypeptides.
CC For example, (I) may be used to treat disorders by rectifying mutations
CC or deletions in a patient's genome that affect the activity of
CC polypeptides by expressing inactive proteins or to supplement the
CC patients own production of polypeptide. Additionally, (I) and its
CC complementary sequences may also be used as DNA probes in diagnostic
CC assays to detect and quantitate the presence of similar nucleic acids
CC in samples, and therefore which patients may be in need of restorative
CC therapy. The polypeptides encoded by (I) may be used as antigens in the
CC production of antibodies specific for polymorphic polypeptides. The
CC antibodies may also be used to down regulate expression and activity.
CC The antibodies may also be used as diagnostic agents for detecting the
CC presence of polymorphic polypeptides in samples.
XX
SQ Sequence 51 BP; 13 A; 14 C; 12 G; 12 T; 0 other;

Query Match 71.0%; Score 14.2; DB 22; Length 51;
Best Local Similarity 84.2%; Pred. No. 1.7e+03;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 CTGTGAGCTGGAAGTCATC 20
|||||||| ||||||| |
Db 25 CTGTGAGCCAGAAGTCAGC 43

RESULT 11
AAI77188
ID AAI77188 standard; DNA; 51 BP.
XX
AC AAI77188;
XX
DT 09-NOV-2001 (first entry)
XX
DE Human silent SNP containing nucleic acid SEQ:4129.
XX
KW Human; single nucleotide polymorphism; SNP; genome; gene therapy;
KW protein therapy; vaccine; probe; diagnostic assay; detection;
KW quantitation; restorative therapy; polymorphic; ds.
XX
OS Homo sapiens.
XX
PN WO200140521-A2.
XX
PD 07-JUN-2001.
XX
PF 30-NOV-2000; 2000WO-US32758.
XX
PR 30-NOV-1999; 99US-0168138.
PR 29-NOV-2000; 2000US-0726173.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shimkets RA, Leach M;
XX
DR WPI; 2001-356160/37.
XX
PT Polymorphic nucleic acid sequences, useful in genetic testing and
PT therapy -
XX
PS Claim 1; Page 1774; 2653pp; English.
XX
CC AAI73060 to AAI79867 represent isolated human polymorphic polynucleotide
CC sequences (I), which contain single nucleotide polymorphisms (SNPs).
CC AAM53114 to AAM53329 represent peptides related to human polymorphic
CC polynucleotide sequences. The sequences can be used in gene and protein
CC therapy, and in vaccine production. (I) and the polypeptides encoded by
CC them may be used in the prevention, diagnosis and treatment of diseases
CC associated with inappropriate expression of polymorphic polypeptides.
CC For example, (I) may be used to treat disorders by rectifying mutations
CC or deletions in a patient's genome that affect the activity of
CC polypeptides by expressing inactive proteins or to supplement the
CC patients own production of polypeptide. Additionally, (I) and its
CC complementary sequences may also be used as DNA probes in diagnostic
CC assays to detect and quantitate the presence of similar nucleic acids
CC in samples, and therefore which patients may be in need of restorative
CC therapy. The polypeptides encoded by (I) may be used as antigens in the
CC production of antibodies specific for polymorphic polypeptides. The
CC antibodies may also be used to down regulate expression and activity.
CC The antibodies may also be used as diagnostic agents for detecting the
CC presence of polymorphic polypeptides in samples.
XX
SQ Sequence 51 BP; 8 A; 14 C; 14 G; 15 T; 0 other;

Query Match 71.0%; Score 14.2; DB 22; Length 51;
Best Local Similarity 84.2%; Pred. No. 1.7e+03;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 CTGTGAGCTGGAAGTCATC 20
|||||||| ||||||| |
Db 4 CTGTGAGCCAGAAGTCAGC 22

RESULT 12
AAI77189
ID AAI77189 standard; DNA; 51 BP.
XX
AC AAI77189;
XX
DT 09-NOV-2001 (first entry)
XX
DE Human silent SNP containing nucleic acid SEQ:4130.
XX
KW Human; single nucleotide polymorphism; SNP; genome; gene therapy;
KW protein therapy; vaccine; probe; diagnostic assay; detection;
KW quantitation; restorative therapy; polymorphic; ds.
XX
OS Homo sapiens.
XX
PN WO200140521-A2.
XX
PD 07-JUN-2001.
XX
PF 30-NOV-2000; 2000WO-US32758.
XX
PR 30-NOV-1999; 99US-0168138.
PR 29-NOV-2000; 2000US-0726173.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shimkets RA, Leach M;
XX
DR WPI; 2001-356160/37.
XX
PT Polymorphic nucleic acid sequences, useful in genetic testing and
PT therapy -
XX
PS Claim 1; Page 1774; 2653pp; English.
XX
CC AAI73060 to AAI79867 represent isolated human polymorphic polynucleotide
CC sequences (I), which contain single nucleotide polymorphisms (SNPs).
CC AAM53114 to AAM53329 represent peptides related to human polymorphic
CC polynucleotide sequences. The sequences can be used in gene and protein
CC therapy, and in vaccine production. (I) and the polypeptides encoded by
CC them may be used in the prevention, diagnosis and treatment of diseases
CC associated with inappropriate expression of polymorphic polypeptides.

```
CC   For example, (I) may be used to treat disorders by rectifying mutations
CC   or deletions in a patient's genome that affect the activity of
CC   polypeptides by expressing inactive proteins or to supplement the
CC   patients own production of polypeptide. Additionally, (I) and its
CC   complementary sequences may also be used as DNA probes in diagnostic
CC   assays to detect and quantitate the presence of similar nucleic acids
CC   in samples, and therefore which patients may be in need of restorative
CC   therapy. The polypeptides encoded by (I) may be used as antigens in the
CC   production of antibodies specific for polymorphic polypeptides. The
CC   antibodies may also be used to down regulate expression and activity.
CC   The antibodies may also be used as diagnostic agents for detecting the
CC   presence of polymorphic polypeptides in samples.
XX
SQ   Sequence 51 BP; 9 A; 14 C; 13 G; 15 T; 0 other;

  Query Match             71.0%;  Score 14.2;  DB 22;  Length 51;
  Best Local Similarity   84.2%;  Pred. No. 1.7e+03;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 CTGTGAGCTGGAAGTCATC 20
           ||||||||  ||||||| |
Db       4 CTGTGAGCCAGAAGTCAGC 22


RESULT 13
ABN35100/c
ID   ABN35100 standard; DNA; 60 BP.
XX
AC   ABN35100;
XX
DT   15-JUL-2002  (first entry)
XX
DE   Human spliced transcript detection oligonucleotide SEQ ID NO:7848.
XX
KW   Human; mouse; rat; splice transcript; detection; RNA transcript;
KW   splice variant; transcriptome; oligonucleotide library; ss.
XX
OS   Homo sapiens.
XX
PN   WO200210449-A2.
XX
PD   07-FEB-2002.
XX
PF   20-JUL-2001; 2001WO-IB01903.
XX
PR   28-JUL-2000; 2000US-221607P.
PR   02-MAY-2001; 2001US-287724P.
XX
PA   (COMP-) COMPUGEN INC.
XX
PI   Shoshan A,  Wasserman A,  Mintz E,  Mintz L,  Faigler S;
XX
DR   WPI; 2002-257383/30.
XX
PT   New oligonucleotide libraries comprising oligonucleotides which
PT   selectively hybridize to mRNAs transcribed from a transcription unit of
PT   a genome, useful for detecting tissue-, pathology-, and
PT   developmental-specific genes  -
XX
PS   Example 1; SEQ ID 7848; 47pp; English.
XX
CC   The present invention describes oligonucleotide libraries for detecting
CC   messenger RNAs that populate a (sub-)transcriptome, where the
CC   (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC   transcription units that populate a genome. The library comprises
CC   several oligonucleotides, each capable of hybridising selectively to a
CC   set of messenger RNAs transcribed from a given transcription unit of
CC   the genome, which encodes one or more messenger RNA splice variants.
CC   The oligonucleotide libraries are useful for detecting mRNAs from a
CC   biological sample, in expression profiling studies, in qualitatively or
CC   quantitatively characterising the corresponding transcriptome, and in
CC   detecting RNA transcripts and splice variants of human or animal
CC   transcriptomes. The libraries may also be used as specialised mini
CC   libraries to detect transcripts of a sub-transcriptome under a
CC   particular biological or pathological state, and so allowing the
CC   detection of tissue- and pathology-specific genes such as those genes
CC   only expressed in specific tissue under a specific pathological
CC   condition; to detect developmental specific genes; and to detect RNA
CC   transcripts and splice variants of a transcriptome of a patient suffering
CC   from a particular disorder. ABN27253 to ABN59589 represent
CC   oligonucleotide sequences from rats, humans and mice, which are used in
CC   the exemplification of the present invention.
CC   N.B. The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 60 BP; 15 A; 16 C; 16 G; 13 T; 0 other;

  Query Match             71.0%;  Score 14.2;  DB 24;  Length 60;
  Best Local Similarity   84.2%;  Pred. No. 1.8e+03;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 CCTGTGAGCTGGAAGTCAT 19
           ||  | |||||||||||||
Db      23 CCAATCAGCTGGAAGTCAT 5


RESULT 14
ABN48570/c
ID   ABN48570 standard; DNA; 60 BP.
XX
AC   ABN48570;
XX
DT   15-JUL-2002  (first entry)
XX
DE   Human spliced transcript detection oligonucleotide SEQ ID NO:21318.
XX
KW   Human; mouse; rat; splice transcript; detection; RNA transcript;
KW   splice variant; transcriptome; oligonucleotide library; ss.
XX
OS   Homo sapiens.
XX
PN   WO200210449-A2.
XX
PD   07-FEB-2002.
XX
PF   20-JUL-2001; 2001WO-IB01903.
XX
PR   28-JUL-2000; 2000US-221607P.
PR   02-MAY-2001; 2001US-287724P.
XX
PA   (COMP-) COMPUGEN INC.
XX
PI   Shoshan A,  Wasserman A,  Mintz E,  Mintz L,  Faigler S;
XX
DR   WPI; 2002-257383/30.
XX
PT   New oligonucleotide libraries comprising oligonucleotides which
PT   selectively hybridize to mRNAs transcribed from a transcription unit of
PT   a genome, useful for detecting tissue-, pathology-, and
PT   developmental-specific genes  -
XX
PS   Example 1; SEQ ID 21318; 47pp; English.
XX
CC   The present invention describes oligonucleotide libraries for detecting
CC   messenger RNAs that populate a (sub-)transcriptome, where the
CC   (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC   transcription units that populate a genome. The library comprises
CC   several oligonucleotides, each capable of hybridising selectively to a
CC   set of messenger RNAs transcribed from a given transcription unit of
CC   the genome, which encodes one or more messenger RNA splice variants.
CC   The oligonucleotide libraries are useful for detecting mRNAs from a
CC   biological sample, in expression profiling studies, in qualitatively or
CC   quantitatively characterising the corresponding transcriptome, and in
CC   detecting RNA transcripts and splice variants of human or animal
CC   transcriptomes. The libraries may also be used as specialised mini
```

CC libraries to detect transcripts of a sub-transcriptome under a
CC particular biological or pathological state, and so allowing the
CC detection of tissue- and pathology-specific genes such as those genes
CC only expressed in specific tissue under a specific pathological
CC condition; to detect developmental specific genes; and to detect RNA
CC transcripts and splice variants of a transcriptome of a patient suffering
CC from a particular disorder. ABN27253 to ABN59589 represent
CC oligonucleotide sequences from rats, humans and mice, which are used in
CC the exemplification of the present invention.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 60 BP; 20 A; 17 C; 13 G; 10 T; 0 other;

Query Match 71.0%; Score 14.2; DB 24; Length 60;
Best Local Similarity 84.2%; Pred. No. 1.8e+03;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 CTGTGAGCTGGAAGTCATC 20
| |||||| ||||||| ||
Db 37 CAGTGAGCAGGAAGTCCTC 19

RESULT 15
ABN52112/c
ID ABN52112 standard; DNA; 65 BP.
XX
AC ABN52112;
XX
DT 15-JUL-2002 (first entry)
XX
DE Mouse spliced transcript detection oligonucleotide SEQ ID NO:24860.
XX
KW Human; mouse; rat; splice transcript; detection; RNA transcript;
KW splice variant; transcriptome; oligonucleotide library; ss.
XX
OS Mus musculus.
XX
PN WO200210449-A2.
XX
PD 07-FEB-2002.
XX
PF 20-JUL-2001; 2001WO-IB01903.
XX
PR 28-JUL-2000; 2000US-221607P.
PR 02-MAY-2001; 2001US-287724P.
XX
PA (COMP-) COMPUGEN INC.
XX
PI Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
XX
DR WPI; 2002-257383/30.
XX
PT New oligonucleotide libraries comprising oligonucleotides which
PT selectively hybridize to mRNAs transcribed from a transcription unit of
PT a genome, useful for detecting tissue-, pathology-, and
PT developmental-specific genes -
XX
PS Example 1; SEQ ID 24860; 47pp; English.
XX
CC The present invention describes oligonucleotide libraries for detecting
CC messenger RNAs that populate a (sub-)transcriptome, where the
CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC transcription units that populate a genome. The library comprises
CC several oligonucleotides, each capable of hybridising selectively to a
CC set of messenger RNAs transcribed from a given transcription unit of
CC the genome, which encodes one or more messenger RNA splice variants.
CC The oligonucleotide libraries are useful for detecting mRNAs from a
CC biological sample, in expression profiling studies, in qualitatively or
CC quantitatively characterising the corresponding transcriptome, and in
CC detecting RNA transcripts and splice variants of human or animal
CC transcriptomes. The libraries may also be used as specialised mini

CC libraries to detect transcripts of a sub-transcriptome under a
CC particular biological or pathological state, and so allowing the
CC detection of tissue- and pathology-specific genes such as those genes
CC only expressed in specific tissue under a specific pathological
CC condition; to detect developmental specific genes; and to detect RNA
CC transcripts and splice variants of a transcriptome of a patient suffering
CC from a particular disorder. ABN27253 to ABN59589 represent
CC oligonucleotide sequences from rats, humans and mice, which are used in
CC the exemplification of the present invention.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 65 BP; 16 A; 17 C; 16 G; 16 T; 0 other;

Query Match 71.0%; Score 14.2; DB 24; Length 65;
Best Local Similarity 84.2%; Pred. No. 1.8e+03;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 CTGTGAGCTGGAAGTCATC 20
||| |||||||||| ||||
Db 46 CTGGAAGCTGGAAGGCATC 28

Search completed: November 23, 2002, 07:03:54
Job time : 97.1 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 06:42:25 ; Search time 16.8 Seconds
(without alignments)
450.869 Million cell updates/sec

Title: US-09-296-264-26
Perfect score: 20
Sequence: 1 cctgtgagctggaagtcatc 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 335578 seqs, 189365133 residues

Total number of hits satisfying chosen parameters: 195614

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	14.4	72.0	77	10	US-09-864-761-25651	Sequence 25651, A
c	2	13.8	69.0	50	12	US-10-036-342-33	Sequence 33, Appl
c	3	13.4	67.0	20	10	US-09-800-631-120	Sequence 120, App
c	4	13.4	67.0	24	10	US-09-906-549-7	Sequence 7, Appli
c	5	12.8	64.0	20	10	US-09-855-722-16	Sequence 16, Appl
	6	12.6	63.0	22	10	US-09-844-268-10	Sequence 10, Appl
	7	12.6	63.0	22	10	US-09-844-705-10	Sequence 10, Appl
	8	12.6	63.0	98	10	US-09-864-761-18752	Sequence 18752, A
	9	12.6	63.0	98	10	US-09-864-761-21771	Sequence 21771, A
	10	12.2	61.0	28	10	US-09-809-545A-79	Sequence 79, Appl
	11	12.2	61.0	36	9	US-09-828-523A-38	Sequence 38, Appl
c	12	12.2	61.0	83	10	US-09-864-761-23154	Sequence 23154, A
	13	12.2	61.0	97	10	US-09-815-343-241	Sequence 241, App
	14	12	60.0	30	10	US-09-972-694-3	Sequence 3, Appli
	15	12	60.0	81	10	US-09-864-761-24915	Sequence 24915, A
	16	12	60.0	86	10	US-09-864-761-25662	Sequence 25662, A
c	17	12	60.0	88	10	US-09-864-761-31535	Sequence 31535, A
c	18	11.8	59.0	28	10	US-09-738-396-36	Sequence 36, Appl
	19	11.8	59.0	37	9	US-10-003-671A-26	Sequence 26, Appl
	20	11.8	59.0	57	10	US-09-955-649-8	Sequence 8, Appli
	21	11.8	59.0	57	10	US-09-973-013-8	Sequence 8, Appli
	22	11.8	59.0	68	9	US-09-466-035-48	Sequence 48, Appl
	23	11.8	59.0	68	10	US-09-912-679-48	Sequence 48, Appl
c	24	11.8	59.0	74	10	US-09-965-602-2	Sequence 2, Appli
c	25	11.8	59.0	74	10	US-09-965-602-14	Sequence 14, Appl
c	26	11.8	59.0	74	10	US-09-965-602-26	Sequence 26, Appl
c	27	11.8	59.0	74	10	US-09-965-602-36	Sequence 36, Appl
c	28	11.8	59.0	74	10	US-09-965-602-50	Sequence 50, Appl
c	29	11.8	59.0	74	10	US-09-965-602-60	Sequence 60, Appl
c	30	11.8	59.0	96	10	US-09-864-761-28861	Sequence 28861, A
c	31	11.6	58.0	26	10	US-09-794-927-49	Sequence 49, Appl
c	32	11.6	58.0	26	10	US-09-795-847-49	Sequence 49, Appl
c	33	11.6	58.0	26	10	US-09-794-743-49	Sequence 49, Appl
c	34	11.6	58.0	26	10	US-09-794-748-49	Sequence 49, Appl
c	35	11.6	58.0	26	10	US-09-794-925-49	Sequence 49, Appl
c	36	11.6	58.0	26	10	US-09-681-442-49	Sequence 49, Appl
c	37	11.6	58.0	27	10	US-09-995-542-22	Sequence 22, Appl
c	38	11.6	58.0	42	10	US-09-147-142-9	Sequence 9, Appli
	39	11.6	58.0	42	10	US-09-147-142-10	Sequence 10, Appl
c	40	11.6	58.0	45	10	US-09-147-142-11	Sequence 11, Appl
	41	11.6	58.0	45	10	US-09-147-142-12	Sequence 12, Appl
c	42	11.6	58.0	48	10	US-09-147-142-13	Sequence 13, Appl
	43	11.6	58.0	48	10	US-09-147-142-14	Sequence 14, Appl
c	44	11.6	58.0	51	10	US-09-794-927-48	Sequence 48, Appl
c	45	11.6	58.0	51	10	US-09-795-847-48	Sequence 48, Appl

ALIGNMENTS

RESULT 1
US-09-864-761-25651/c
; Sequence 25651, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 25651
; LENGTH: 77
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC021735.4
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.8
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.9
; OTHER INFORMATION: EST_HUMAN HIT: AA766755.1, EVALUE 2.00e-20
; OTHER INFORMATION: NT HIT: AF114156.1, EVALUE 1.00e-14
US-09-864-761-25651

Query Match 72.0%; Score 14.4; DB 10; Length 77;
Best Local Similarity 93.8%; Pred. No. 2e+02;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy       2 CTGTGAGCTGGAAGTC 17
           |||||| |||||||||
Db      77 CTGTGACCTGGAAGTC 62
```

RESULT 2
US-10-036-342-33/c
; Sequence 33, Application US/10036342
; Patent No. US20020090681A1
; GENERAL INFORMATION:
; APPLICANT: Desnoyers,Luc
; APPLICANT: Eaton,Dan L.
; APPLICANT: Goddard,Audrey
; APPLICANT: Godowski,Paul J.
; APPLICANT: Gurney,Austin L.
; APPLICANT: Pan,James
; APPLICANT: Stewart,Timothy A.
; APPLICANT: Watanabe,Colin K.
; APPLICANT: Wood,William I.
; APPLICANT: Zhang,Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3030R1C5
; CURRENT APPLICATION NUMBER: US/10/036,342
; CURRENT FILING DATE: 2001-12-26
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/112514
; PRIOR FILING DATE: 1998-12-15
; PRIOR APPLICATION NUMBER: 60/113300
; PRIOR FILING DATE: 1998-12-22
; PRIOR APPLICATION NUMBER: 60/113430
; PRIOR FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: 60/113605
; PRIOR FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: 60/113621
; PRIOR FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: 60/114140
; PRIOR FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: 60/115552
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/116843
; PRIOR FILING DATE: 1999-01-22
; PRIOR APPLICATION NUMBER: 60/125774
; PRIOR FILING DATE: 1999-03-23
; PRIOR APPLICATION NUMBER: 60/125778
; PRIOR FILING DATE: 1999-03-23
; PRIOR APPLICATION NUMBER: 60/125826
; PRIOR FILING DATE: 1999-03-24
; PRIOR APPLICATION NUMBER: 60/127035
; PRIOR FILING DATE: 1999-03-31
; PRIOR APPLICATION NUMBER: 60/127706
; PRIOR FILING DATE: 1999-04-05
; PRIOR APPLICATION NUMBER: 60/129122
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: 60/130359
; PRIOR FILING DATE: 1999-04-21
; PRIOR APPLICATION NUMBER: 60/131270
; PRIOR FILING DATE: 1999-04-27
; PRIOR APPLICATION NUMBER: 60/131272
; PRIOR FILING DATE: 1999-04-27
; PRIOR APPLICATION NUMBER: 60/131291
; PRIOR FILING DATE: 1999-04-27
; PRIOR APPLICATION NUMBER: 60/132371
; PRIOR FILING DATE: 1999-05-04
; PRIOR APPLICATION NUMBER: 60/132379
; PRIOR FILING DATE: 1999-05-04
; PRIOR APPLICATION NUMBER: 60/132383
; PRIOR FILING DATE: 1999-05-04
; PRIOR APPLICATION NUMBER: 60/135750
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: 60/138166
; PRIOR FILING DATE: 1999-06-08
; PRIOR APPLICATION NUMBER: 60/144791
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 60/146970
; PRIOR FILING DATE: 1999-08-03
; PRIOR APPLICATION NUMBER: 60/162506
; PRIOR FILING DATE: 1999-10-29
; PRIOR APPLICATION NUMBER: 09/311832
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: 09/380142
; PRIOR FILING DATE: 1999-08-25
; PRIOR APPLICATION NUMBER: 09/644848
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 09/747259
; PRIOR FILING DATE: 2000-12-20
; PRIOR APPLICATION NUMBER: 09/816744
; PRIOR FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: 09/854208
; PRIOR FILING DATE: 2001-05-10
; PRIOR APPLICATION NUMBER: 09/854280
; PRIOR FILING DATE: 2001-05-10
; PRIOR APPLICATION NUMBER: 09/874503
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: 09/869599
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: 09/908,827
; PRIOR FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: PCT/US99/10733
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: PCT/US99/28551
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30720
; PRIOR FILING DATE: 1999-12-22
; PRIOR APPLICATION NUMBER: PCT/US00/05601
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: PCT/US00/05841
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: PCT/US00/14042
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: PCT/US00/15264
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: PCT/US00/23522
; PRIOR FILING DATE: 2000-08-23
; PRIOR APPLICATION NUMBER: PCT/US00/23328
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: PCT/US00/32678
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: PCT/US00/34956

```
;  PRIOR FILING DATE: 2000-12-20
;  PRIOR APPLICATION NUMBER: PCT/US01/06520
;  PRIOR FILING DATE: 2001-02-28
;  PRIOR APPLICATION NUMBER: PCT/US01/17800
;  PRIOR FILING DATE: 2001-06-01
;  PRIOR APPLICATION NUMBER: PCT/US01/19692
;  PRIOR FILING DATE: 2001-06-20
;  PRIOR APPLICATION NUMBER: PCT/US01/21066
;  PRIOR FILING DATE: 2001-06-29
;  PRIOR APPLICATION NUMBER: PCT/US01/21735
;  PRIOR FILING DATE: 2001-07-09
;  NUMBER OF SEQ ID NOS: 80
; SEQ ID NO 33
;   LENGTH: 50
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Synthetic oligonucleotide probe
US-10-036-342-33

  Query Match             69.0%;  Score 13.8;  DB 12;  Length 50;
  Best Local Similarity   88.2%;  Pred. No. 3.7e+02;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      2 CTGTGAGCTGGAAGTCA 18
          ||||||| ||| |||||
Db     50 CTGTGAGTTGGCAGTCA 34


RESULT 3
US-09-800-631-120/c
; Sequence 120, Application US/09800631
; Patent No. US20020082228A1
; GENERAL INFORMATION:
;  APPLICANT: Hong Zhang
;  APPLICANT:  Jacqueline Wyatt
;  TITLE OF INVENTION: ANTISENSE MODULATION OF BH3 INTERACTING DOMAIN DEATH AGONIST EXPR
;  FILE REFERENCE: ISPH-0544
;  CURRENT APPLICATION NUMBER: US/09/800,631
;  CURRENT FILING DATE:  2001-03-07
;  PRIOR APPLICATION NUMBER: US/09/657,346
;  PRIOR FILING DATE: 2000-09-07
;  NUMBER OF SEQ ID NOS: 175
; SEQ ID NO 120
;   LENGTH: 20
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Antisense Oligonucleotide
US-09-800-631-120

  Query Match             67.0%;  Score 13.4;  DB 10;  Length 20;
  Best Local Similarity   93.3%;  Pred. No. 5e+02;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      2 CTGTGAGCTGGAAGT 16
          ||||||||||| |||
Db     20 CTGTGAGCTGGCAGT 6


RESULT 4
US-09-906-549-7/c
; Sequence 7, Application US/09906549
; Patent No. US20020099193A1
; GENERAL INFORMATION:
;  APPLICANT: Kaeppler, Shawn
;  APPLICANT:  Springer, Nathan
;  APPLICANT:  Phillips, Ronald
;  TITLE OF INVENTION: Polycomb Gene from Maize - ZMFIE2
;  FILE REFERENCE: WARF101US
;  CURRENT APPLICATION NUMBER: US/09/906,549
;  CURRENT FILING DATE:  2001-07-16
;  NUMBER OF SEQ ID NOS: 11
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
;   LENGTH: 24
;   TYPE: DNA
;   ORGANISM: Zea mays
US-09-906-549-7

  Query Match             67.0%;  Score 13.4;  DB 10;  Length 24;
  Best Local Similarity   93.3%;  Pred. No. 5.2e+02;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      6 GAGCTGGAAGTCATC 20
          |||||||| ||||||
Db     24 GAGCTGGAGGTCATC 10


RESULT 5
US-09-855-722-16/c
; Sequence 16, Application US/09855722
; Patent No. US20020049306A1
; GENERAL INFORMATION:
;  APPLICANT: Sakano, Seiji
;  APPLICANT:  Itoh, Akira
;  TITLE OF INVENTION: DIFFERENTIATION-SUPPRESSIVE POLYPEPTIDE
;  FILE REFERENCE: KP-8576
;  CURRENT APPLICATION NUMBER: US/09/855,722
;  CURRENT FILING DATE:  2001-05-16
;  PRIOR APPLICATION NUMBER: 09/214,278
;  PRIOR FILING DATE: 1999-01-26
;  NUMBER OF SEQ ID NOS: 32
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 16
;   LENGTH: 20
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: synthetic DNA
US-09-855-722-16

  Query Match             64.0%;  Score 12.8;  DB 10;  Length 20;
  Best Local Similarity   87.5%;  Pred. No. 1e+03;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      2 CTGTGAGCTGGAAGTC 17
          ||||||| |||| |||
Db     20 CTGTGAGGTGGATGTC 5


RESULT 6
US-09-844-268-10
; Sequence 10, Application US/09844268
; Patent No. US20020129395A1
; GENERAL INFORMATION:
;  APPLICANT: BOSWORTH, BRAD
;  APPLICANT:  VOGELI, PETER
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS TO IDENTIFY SWINE GENETICALLY
;  TITLE OF INVENTION:  RESISTANT TO F18 E. COLI ASSOCIATED DISEASES
;  FILE REFERENCE: 21419/90368
;  CURRENT APPLICATION NUMBER: US/09/844,268
;  CURRENT FILING DATE:  2001-04-27
;  PRIOR APPLICATION NUMBER: 09/443,766
;  PRIOR FILING DATE: 1999-11-19
;  NUMBER OF SEQ ID NOS: 13
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
;   LENGTH: 22
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Primer
US-09-844-268-10

  Query Match             63.0%;  Score 12.6;  DB 10;  Length 22;
```

```
  Best Local Similarity   78.9%;  Pred. No. 1.3e+03;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 CCTGTGAGCTGGAAGTCAT 19
           ||||||   ||||||| ||
Db       4 CCTGTGCCTTGGAAGTGAT 22


RESULT 7
US-09-844-705-10
; Sequence 10, Application US/09844705
; Patent No. US20020133836A1
; GENERAL INFORMATION:
;  APPLICANT: BOSWORTH, BRAD
;  APPLICANT:  VOGELI, PETER
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS TO IDENTIFY SWINE GENETICALLY
;  TITLE OF INVENTION:  RESISTANT TO F18 E. COLI ASSOCIATED DISEASES
;  FILE REFERENCE: 21419/90368
;  CURRENT APPLICATION NUMBER: US/09/844,705
;  CURRENT FILING DATE:  2001-04-27
;  PRIOR APPLICATION NUMBER: 09/443,766
;  PRIOR FILING DATE: 1999-11-19
;  NUMBER OF SEQ ID NOS: 13
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
;   LENGTH: 22
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Primer
US-09-844-705-10

  Query Match             63.0%;  Score 12.6;  DB 10;  Length 22;
  Best Local Similarity   78.9%;  Pred. No. 1.3e+03;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 CCTGTGAGCTGGAAGTCAT 19
           ||||||   ||||||| ||
Db       4 CCTGTGCCTTGGAAGTGAT 22


RESULT 8
US-09-864-761-18752
; Sequence 18752, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 18752
;   LENGTH: 98
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AL033539.16
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.7
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.7
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.2
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.3
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3.2
;   OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.4
;   OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL =   3
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.5
;   OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.7
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.6
;   OTHER INFORMATION: NT HIT: AB000489.1, EVALUE 8.20e-01
;   OTHER INFORMATION: EST_HUMAN HIT: BF681239.1, EVALUE 1.20e+00
US-09-864-761-18752

  Query Match             63.0%;  Score 12.6;  DB 10;  Length 98;
  Best Local Similarity   78.9%;  Pred. No. 1.6e+03;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 CTGTGAGCTGGAAGTCATC 20
           || |||  ||| |||||||
Db      10 CTATGAAATGGCAGTCATC 28


RESULT 9
US-09-864-761-21771
; Sequence 21771, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
```

```
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 21771
;   LENGTH: 98
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AL033539.17
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.1
;   OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 0.99
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.2
;   OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL =   1
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.1
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.98
;   OTHER INFORMATION: EST_HUMAN HIT: BF681239.1, EVALUE 1.20e+00
;   OTHER INFORMATION: NT HIT: AB000489.1, EVALUE 8.20e-01
US-09-864-761-21771

  Query Match             63.0%;  Score 12.6;  DB 10;  Length 98;
  Best Local Similarity   78.9%;  Pred. No. 1.6e+03;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 CTGTGAGCTGGAAGTCATC 20
           || |||  ||| |||||||
Db      10 CTATGAAATGGCAGTCATC 28


RESULT 10
US-09-809-545A-79
; Sequence 79, Application US/09809545A
; Patent No. US20020110804A1
; GENERAL INFORMATION:
;  APPLICANT: Stanton, Lawrence W.
;  APPLICANT:  White, R. Tyler
;  TITLE OF INVENTION: SECRETED FACTORS
;  FILE REFERENCE: SCIOS.017A
;  CURRENT APPLICATION NUMBER: US/09/809,545A
;  CURRENT FILING DATE:  2001-03-14
;  NUMBER OF SEQ ID NOS: 84
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 79
;   LENGTH: 28
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Oligos corresponding to polylinker sequence.
US-09-809-545A-79

  Query Match             61.0%;  Score 12.2;  DB 10;  Length 28;
  Best Local Similarity   82.4%;  Pred. No. 2.1e+03;
  Matches   14;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 CCTGTGAGCTGGAAGTC 17
           || ||| ||||||| ||
Db       5 CCAGTGTGCTGGAATTC 21


RESULT 11
US-09-828-523A-38
; Sequence 38, Application US/09828523A
; Patent No. US20020168697A1
; GENERAL INFORMATION:
;  APPLICANT: The Pharmacia & Upjohn Company
;  TITLE OF INVENTION: ANTIMICROBIAL METHODS AND MATERIALS
;  FILE REFERENCE: 268.62120101
;  CURRENT APPLICATION NUMBER: US/09/828,523A
;  CURRENT FILING DATE:  2001-04-06
;  PRIOR APPLICATION NUMBER: 60/266,327
;  PRIOR FILING DATE: 2000-04-06
;  NUMBER OF SEQ ID NOS: 99
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 38
;   LENGTH: 36
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Oligonucleotide Primer.
US-09-828-523A-38

  Query Match             61.0%;  Score 12.2;  DB 9;  Length 36;
  Best Local Similarity   82.4%;  Pred. No. 2.2e+03;
  Matches   14;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       3 TGTGAGCTGGAAGTCAT 19
           | |||||| | ||||||
Db       4 TATGAGCTCGCAGTCAT 20


RESULT 12
US-09-864-761-23154/c
; Sequence 23154, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
```

```
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 23154
;   LENGTH: 83
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AC004969.1
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.4
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.4
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.3
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.5
;   OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.3
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.3
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.4
;   OTHER INFORMATION: NT HIT: AB035414.1, EVALUE 4.30e-02
;   OTHER INFORMATION: EST_HUMAN HIT: AI221371.1, EVALUE 8.00e-20
US-09-864-761-23154

  Query Match             61.0%;  Score 12.2;  DB 10;  Length 83;
  Best Local Similarity   82.4%;  Pred. No. 2.5e+03;
  Matches   14;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 CTGTGAGCTGGAAGTCA 18
           ||| ||||||| || ||
Db      83 CTGGGAGCTGGGAGACA 67


RESULT 13
US-09-815-343-241
; Sequence 241, Application US/09815343
; Patent No. US20010055596A1
; GENERAL INFORMATION:
;  APPLICANT: Meagher, Madeleine
;  APPLICANT:  Xu, Jiangchun
;  APPLICANT:  King, Gordon E.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
;  TITLE OF INVENTION:  DIAGNOSIS OF COLON CANCER
;  FILE REFERENCE: 210121.504
;  CURRENT APPLICATION NUMBER: US/09/815,343
;  CURRENT FILING DATE:  2001-03-22
;  NUMBER OF SEQ ID NOS: 1556
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 241
;   LENGTH: 97
;   TYPE: DNA
;   ORGANISM: Homo sapien
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(97)
;   OTHER INFORMATION: n = A,T,C or G
US-09-815-343-241

  Query Match             61.0%;  Score 12.2;  DB 10;  Length 97;
  Best Local Similarity   82.4%;  Pred. No. 2.6e+03;
  Matches   14;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 CCTGTGAGCTGGAAGTC 17
           || ||| ||||||| ||
Db      25 CCAGTGTGCTGGAATTC 41


RESULT 14
US-09-972-694-3
; Sequence 3, Application US/09972694
; Patent No. US20020102587A1
; GENERAL INFORMATION:
;  APPLICANT: DELANEY, ALLEN
;  APPLICANT:  YOGANATHAN, THILLAINATHAN
;  TITLE OF INVENTION: G PROTEIN COUPLED RECEPTOR KINASE 5
;  TITLE OF INVENTION:  (GRK5) AND ITS USES
;  FILE REFERENCE: KINE027CON
;  CURRENT APPLICATION NUMBER: US/09/972,694
;  CURRENT FILING DATE:  2001-10-04
;  PRIOR APPLICATION NUMBER: PCT/US01/21479
;  PRIOR FILING DATE: 2001-09-20
;  PRIOR APPLICATION NUMBER: 60/237,423
;  PRIOR FILING DATE: 2000-10-02
;  NUMBER OF SEQ ID NOS: 4
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
;   LENGTH: 30
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-972-694-3

  Query Match             60.0%;  Score 12;  DB 10;  Length 30;
  Best Local Similarity   75.0%;  Pred. No. 2.7e+03;
  Matches   15;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1 CCTGTGAGCTGGAAGTCATC 20
           ||   |||||||||  ||||
Db       5 CCATGGAGCTGGAAAACATC 24


RESULT 15
US-09-864-761-24915
; Sequence 24915, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
```

```
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 24915
;   LENGTH: 81
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AC016663.2
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3.4
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.8
;   OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.7
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.7
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 8.1
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.5
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3.4
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.8
;   OTHER INFORMATION: NT HIT: AB026898.1, EVALUE 3.00e-15
;   OTHER INFORMATION: SWISSPROT HIT: P13508, EVALUE 6.50e-01
;   OTHER INFORMATION: EST_HUMAN HIT: BF591249.1, EVALUE 2.00e-14
US-09-864-761-24915

  Query Match             60.0%;  Score 12;  DB 10;  Length 81;
  Best Local Similarity   75.0%;  Pred. No. 3.2e+03;
  Matches   15;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy       1 CCTGTGAGCTGGAAGTCATC 20
           || | | | |||||||||  |
Db      23 CCGGAGGGGTGGAAGTCAGC 42


Search completed: November 23, 2002, 07:10:42
Job time : 17.8 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 26, 2002, 04:16:10 ; Search time 798.5 Seconds
(without alignments)
405.648 Million cell updates/sec

Title: US-09-296-264-26
Perfect score: 20
Sequence: 1 cctgtgagctggaagtcatc 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 357874

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	18	90.0	37	9	AA048244	AA048244 mj27a04.r
c	2	15.2	76.0	65	9	AA145279	AA145279 ms10c10.r
	3	14.8	74.0	37	9	AA102976	AA102976 mo10d11.r
c	4	14.8	74.0	83	9	AA414440	AA414440 vd08g10.s
c	5	14.4	72.0	79	9	AA787956	AA787956 r3d09a1.f
c	6	14.2	71.0	50	9	AU104488	AU104488 AU104488
c	7	14.2	71.0	68	9	AI561510	AI561510 vw92b09.x
	8	14.2	71.0	76	9	AA451089	AA451089 vf87d06.r
	9	13.8	69.0	65	12	BE957944	BE957944 601654129
c	10	13.8	69.0	68	17	AZ467019	AZ467019 1M0278E10
	11	13.8	69.0	79	14	N67141	N67141 yz46h01.s1
c	12	13.8	69.0	95	14	BQ756524	BQ756524 EBem09_SQ
	13	13.6	68.0	67	14	R72481	R72481 yj90f05.r1
c	14	13.4	67.0	52	10	BE318017	BE318017 NF061E07L
	15	13.4	67.0	76	14	D38680	D38680 HUMCI215 Hu
	16	13.4	67.0	80	9	AI437218	AI437218 fb39b01.y
	17	13.2	66.0	35	17	AZ479885	AZ479885 1M0300L21
c	18	13.2	66.0	64	13	BI142469	BI142469 SWOv3MCAM
	19	13.2	66.0	66	10	BE291121	BE291121 601086604
c	20	13.2	66.0	66	14	H55131	H55131 CHR220070 C
c	21	13.2	66.0	69	14	T12642	T12642 CHR90162 Ch
c	22	13.2	66.0	73	10	BE420479	BE420479 SWOvL2CAS
	23	13.2	66.0	92	13	BI853339	BI853339 603379919
c	24	13.2	66.0	97	9	AA512842	AA512842 ne85c08.s
	25	13.2	66.0	100	12	BF181495	BF181495 601805686
	26	13	65.0	54	17	BH638209	BH638209 1008021A0
c	27	12.8	64.0	30	17	TA21H04Q	AL453803 T. brucei
	28	12.8	64.0	42	17	AZ403048	AZ403048 1M0170I10
	29	12.8	64.0	50	9	AU103373	AU103373 AU103373
	30	12.8	64.0	57	14	T12567	T12567 CHR90087 Ch
c	31	12.8	64.0	58	14	T61500	T61500 yb73c02.s1
	32	12.8	64.0	64	14	BQ549177	BQ549177 ik89h01.y
	33	12.8	64.0	65	10	AW059697	AW059697 AHuTH.bss
c	34	12.8	64.0	67	14	T91872	T91872 yd54d03.s1
c	35	12.8	64.0	69	14	T59369	T59369 yb57h04.s1
c	36	12.8	64.0	69	17	BH216864	BH216864 1006047G0
c	37	12.8	64.0	70	9	AA553179	AA553179 vl13e01.r
c	38	12.8	64.0	76	9	AA517771	AA517771 vh94a09.r
	39	12.8	64.0	79	9	AA670839	AA670839 vs90e06.r
c	40	12.8	64.0	86	9	AI250833	AI250833 qx04g08.x
c	41	12.8	64.0	94	10	AW827260	AW827260 hl97f08.x
c	42	12.8	64.0	98	17	AL768126	AL768126 Arabidops
c	43	12.8	64.0	100	9	AA855128	AA855128 aj78a01.s
	44	12.8	64.0	100	14	BQ331911	BQ331911 MR4-ET013
c	45	12.6	63.0	33	14	T64728	T64728 yc48e07.r1

ALIGNMENTS

RESULT 1
AA048244/c
LOCUS AA048244 37 bp mRNA linear EST 09-SEP-1996
DEFINITION mj27a04.r1 Soares mouse embryo NbME13.5 14.5 Mus musculus cDNA clone IMAGE:477294 5' similar to SW:A5_XENLA P28824 A5 PROTEIN PRECURSOR ;, mRNA sequence.
ACCESSION AA048244
VERSION AA048244.1 GI:1527914
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 37)
AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.
TITLE The WashU-HHMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Marra M/Mouse EST Project
WashU-HHMI Mouse EST Project
Washington University School of MedicineP
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the

```
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:288038
            Trace considered overall poor quality
            Possible reversed clone: similarity on wrong strand
            Seq primer: -28M13 rev2 from Amersham
            High quality sequence stop: 1.
FEATURES              Location/Qualifiers
     source           1. .37
                      /organism="Mus musculus"
                      /strain="C57BL/6J"
                      /db_xref="taxon:10090"
                      /clone="IMAGE:477294"
                      /clone_lib="Soares mouse embryo NbME13.5 14.5"
                      /sex="unknown"
                      /tissue_type="embryo"
                      /dev_stage="13.5-14.5dpc total fetus"
                      /lab_host="DH10B"
                      /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                      polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                      was primed with a Not I - oligo(dT) primer [5'
                      TGTTACCAATCTGAAGTGGGAGCGGCCGCGGAAATTTTTTTTTTTTTTTTTTTTTT
                      T 3'], on equal amounts of mRNA from 2 13.5dpc and 2
                      14.5dpc embryos [total RNA provided by  Minoru Ko, Wayne
                      State Univ., from 2 ]; double-stranded cDNA was ligated to
                      Eco RI adaptors (Pharmacia), digested with Not I and
                      cloned into the Not I and Eco RI sites of the modified
                      pT7T3 vector. Library went through one round of
                      normalization, and was constructed by Bento Soares and
                      M.Fatima Bonaldo.  "
BASE COUNT        9 a     12 c      9 g      7 t
ORIGIN

  Query Match             90.0%;  Score 18;  DB 9;  Length 37;
  Best Local Similarity   100.0%;  Pred. No. 1.3e+02;
  Matches   18;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3 TGTGAGCTGGAAGTCATC 20
           ||||||||||||||||||
Db      37 TGTGAGCTGGAAGTCATC 20


RESULT 2
AA145279/c
LOCUS       AA145279                 65 bp    mRNA    linear   EST 12-FEB-1997
DEFINITION  ms10c10.r1 Stratagene mouse skin (#937313) Mus musculus cDNA clone
            IMAGE:606546 5', mRNA sequence.
ACCESSION   AA145279
VERSION     AA145279.1  GI:1714652
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 65)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:371978
            Seq primer: -28m13 rev1 ET from Amersham.
FEATURES              Location/Qualifiers
     source           1. .65
                      /organism="Mus musculus"
                      /strain="C57BL/6"
                      /db_xref="taxon:10090"
                      /clone="IMAGE:606546"
                      /clone_lib="Stratagene mouse skin (#937313)"
                      /sex="females"
                      /tissue_type="whole skin"
                      /dev_stage="11 weeks old"
                      /lab_host="SOLR (kanamycin resistant)"
                      /note="Organ: skin; Vector: pBluescript SK-; Site_1: EcoRI
                      ; Site_2: XhoI; Cloned unidirectionally.  Primer: Oligo
                      dT. Whole skin from 11 week old C57BL/6 female mice.
                      Average insert size: 1.0 kb; Uni-ZAP XR Vector; ~5'
                      adaptor sequence: 5' GAATTCGGCACGAG 3' ~3' adaptor
                      sequence: 5' CTCGAGTTTTTTTTTTTTTTTT 3'"
BASE COUNT       17 a     15 c     17 g     16 t
ORIGIN

  Query Match             76.0%;  Score 15.2;  DB 9;  Length 65;
  Best Local Similarity   85.0%;  Pred. No. 3.6e+03;
  Matches   17;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 CCTGTGAGCTGGAAGTCATC 20
           ||||  |||||||||| |||
Db      20 CCTGCAAGCTGGAAGTTATC 1


RESULT 3
AA102976
LOCUS       AA102976                 37 bp    mRNA    linear   EST 29-OCT-1996
DEFINITION  mo10d11.r1 Life Tech mouse embryo 10 5dpc 10665016 Mus musculus
            cDNA clone IMAGE:553173 5' similar to TR:G1236243 G1236243
            STEROIDOGENIC ACUTE REGULATORY PROTEIN. ;, mRNA sequence.
ACCESSION   AA102976
VERSION     AA102976.1  GI:1649137
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 37)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:333965
            Trace considered overall poor quality
            Possible reversed clone: similarity on wrong strand
            Seq primer: -28M13 rev1 from Amersham
            High quality sequence stop: 1.
FEATURES              Location/Qualifiers
     source           1. .37
                      /organism="Mus musculus"
                      /strain="C57BL/6J"
                      /db_xref="taxon:10090"
                      /clone="IMAGE:553173"
                      /clone_lib="Life Tech mouse embryo 10 5dpc 10665016"
                      /tissue_type="embryo"
                      /dev_stage="10.5dpc embryos"
                      /lab_host="DH10B"
```

/note="Organ: whole embryo; Vector: pCMV-SPORT2; Site_1: SalI; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT. 10.5dpc embryos. pCMV-SPORT2 vector."
BASE COUNT 5 a 10 c 10 g 12 t
ORIGIN

Query Match 74.0%; Score 14.8; DB 9; Length 37;
Best Local Similarity 88.9%; Pred. No. 4.2e+03;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 CTGTGAGCTGGAAGTCAT 19
||||| |||||||||| |
Db 20 CTGTGTGCTGGAAGTCCT 37

RESULT 4
AA414440/c
LOCUS AA414440 83 bp mRNA linear EST 04-AUG-1997
DEFINITION vd08g10.s1 Knowles Solter mouse 2 cell Mus musculus cDNA clone IMAGE:791970 5', mRNA sequence.
ACCESSION AA414440
VERSION AA414440.1 GI:2074577
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 83)
AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.
TITLE The WashU-HHMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Marra M/Mouse EST Project
WashU-HHMI Mouse EST Project
Washington University School of MedicineP
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:476378
Seq primer: -40m13 fwd. ET from Amersham.
FEATURES Location/Qualifiers
source 1. .83
/organism="Mus musculus"
/strain="C57BL/6J x DBA/2J F1"
/db_xref="taxon:10090"
/clone="IMAGE:791970"
/clone_lib="Knowles Solter mouse 2 cell"
/tissue_type="embryo"
/dev_stage="2-cell"
/lab_host="DH10B"
/note="Organ: embryo; Vector: pBluescribe (modified); Site_1: MluI; Site_2: SalI; Cloned unidirectionally from mRNA prepared from 13,500 2-cell stage embryos. Primer: SalI(dT): 5'-CGGTCGACCGTCGACCGTTTTTTTTTTTTTT-3'. cDNAs were cloned into the MluI/SalI sites of a modified pBluescribe vector using commercial linkers (NEB). Average insert size: 1.2 kb."
BASE COUNT 23 a 17 c 26 g 17 t
ORIGIN

Query Match 74.0%; Score 14.8; DB 9; Length 83;
Best Local Similarity 88.9%; Pred. No. 6.2e+03;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 CTGTGAGCTGGAAGTCAT 19
||||| ||||||| ||||
Db 56 CTGTGTGCTGGAATTCAT 39

RESULT 5
AA787956/c
LOCUS AA787956 79 bp mRNA linear EST 28-DEC-1998
DEFINITION r3d09a1.f1 Aspergillus nidulans 24hr asexual developmental and vegetative cDNA lambda zap library Emericella nidulans cDNA clone r3d09a1 3', mRNA sequence.
ACCESSION AA787956
VERSION AA787956.1 GI:4064683
KEYWORDS EST.
SOURCE Emericella nidulans.
ORGANISM Emericella nidulans
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes; Eurotiales; Trichocomaceae; Emericella.
REFERENCE 1 (bases 1 to 79)
AUTHORS Kupfer,D., Gray,J., Hausner,J., Lai,H., Martin,W., Aramayo,R., Prade,R. and Roe,B.
TITLE An Aspergillus nidulans EST Database
JOURNAL Unpublished (1998)
COMMENT On Feb 6, 1998 this sequence version replaced gi:2848187.
Other_ESTs: r3d09a1.r1
Contact: Bruce A. Roe, University of Oklahoma, broe@ou.edu
Department of Chemistry and Biochemistry
Advanced Center for Genome Technology, University of Oklahoma
620 Parrington Oval, Norman, OK 73019, USA
Tel: 405 325 4912
Fax: 405 325 7762
Email: broe@ou.edu
We anticipate the future release of the cDNA clones to the Fungal Genetics Stock Center
Seq primer: M13-20
High quality sequence stop: 349.
FEATURES Location/Qualifiers
source 1. .79
/organism="Emericella nidulans"
/strain="FGSC A26"
/db_xref="taxon:162425"
/clone="r3d09a1"
/clone_lib="Aspergillus nidulans 24hr asexual developmental and vegetative cDNA lambda zap library"
/tissue_type="vegetative mycelia, asexual structures"
/note="Vector: pBlueScript SK-; Site_1: EcoRI; Site_2: XhoI; 5' end of cDNA cloned into EcoRI site of pBluescript 3' end of cDNA cloned into XhoI site of pBluescript"
BASE COUNT 21 a 21 c 15 g 22 t
ORIGIN

Query Match 72.0%; Score 14.4; DB 9; Length 79;
Best Local Similarity 93.8%; Pred. No. 9.3e+03;
Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 5 TGAGCTGGAAGTCATC 20
||||||||||||| |||
Db 49 TGAGCTGGAAGTGATC 34

RESULT 6
AU104488/c
LOCUS AU104488 50 bp mRNA linear EST 30-AUG-2001
DEFINITION AU104488 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone REC00573, mRNA sequence.
ACCESSION AU104488
VERSION AU104488.1 GI:13554009
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 50)
AUTHORS Suzuki,Y., Taira,H., Tsunoda,T., Mizushima-Sugano,J., Sese,J., Hata ,H., Ota,T., Isogai,T., Tanaka,T., Morishita,S., Okubo,K., Sakaki ,Y., Nakamura,Y., Suyama,A. and Sugano,S.

```
  TITLE     Diverse transcriptional initiation revealed by fine, large-scale
            mapping of mRNA start sites
  JOURNAL   EMBO Rep. 2 (5), 388-393 (2001)
  MEDLINE   21270072
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp
            Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano
            ,S. Construction and characterization of a full length-enriched and
            a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="REC00573"
                     /clone_lib="Sugano Homo sapiens cDNA library"
                     /note="Differential display comparison of untreated and
                     dimethylfumarate treated U937 cells"
BASE COUNT       11 a     19 c     11 g      9 t
ORIGIN

  Query Match             71.0%;  Score 14.2;  DB 9;  Length 50;
  Best Local Similarity   84.2%;  Pred. No. 9.3e+03;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 CTGTGAGCTGGAAGTCATC 20
           | ||||||||||| || |
Db      47 CGGTGAGCTGGAAGGCAAC 29


RESULT 7
AI561510/c
LOCUS       AI561510                  68 bp    mRNA    linear   EST 25-MAR-1999
DEFINITION  vw92b09.x1 Stratagene mouse skin (#937313) Mus musculus cDNA clone
            IMAGE:1262393 3', mRNA sequence.
ACCESSION   AI561510
VERSION     AI561510.1  GI:4512855
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 68)
  AUTHORS   Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,
            Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person
            ,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter
            ,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R.,
            Waterston,R. and Wilson,R.
  TITLE     The WashU-NCI Mouse EST Project 1999
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Marra M/WashU-NCI Mouse EST Project 1999
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:664945
            This clone was previously sequenced on the 5' end only, this new
            data is from the 3' end
            High quality sequence stop: 52.
FEATURES             Location/Qualifiers
     source          1. .68
                     /organism="Mus musculus"
                     /strain="C57BL/6"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:1262393"
                     /clone_lib="Stratagene mouse skin (#937313)"
                     /sex="females"
                     /tissue_type="whole skin"
                     /dev_stage="11 weeks old"
                     /lab_host="SOLR (kanamycin resistant)"
                     /note="Organ: skin; Vector: pBluescript SK-; Site_1: EcoRI
                     ; Site_2: XhoI; Cloned unidirectionally.  Primer: Oligo
                     dT. Whole skin from 11 week old C57BL/6 female mice.
                     Average insert size: 1.0 kb; Uni-ZAP XR Vector; ~5'
                     adaptor sequence: 5' GAATTCGGCACGAG 3' ~3' adaptor
                     sequence: 5' CTCGAGTTTTTTTTTTTTTTTT 3'"
BASE COUNT       24 a     19 c      7 g     18 t
ORIGIN

  Query Match             71.0%;  Score 14.2;  DB 9;  Length 68;
  Best Local Similarity   84.2%;  Pred. No. 1.1e+04;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 CTGTGAGCTGGAAGTCATC 20
           ||||||| | || ||||||
Db      57 CTGTGAGTTTGAGGTCATC 39


RESULT 8
AA451089
LOCUS       AA451089                  76 bp    mRNA    linear   EST 04-JUN-1997
DEFINITION  vf87d06.r1 Soares_mammary_gland_NbMMG Mus musculus cDNA clone
            IMAGE:850763 5', mRNA sequence.
ACCESSION   AA451089
VERSION     AA451089.1  GI:2164759
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 76)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:502915
            Seq primer: -28m13 rev2 ET from Amersham
            High quality sequence stop: 67.
FEATURES             Location/Qualifiers
     source          1. .76
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:850763"
                     /clone_lib="Soares_mammary_gland_NbMMG"
                     /sex="male"
                     /tissue_type="mammary gland"
                     /dev_stage="4 weeks"
                     /lab_host="DH10B"
                     /note="Organ: mammary gland; Vector: pT7T3D-Pac (Pharmacia
                     ) with a modified polylinker; Site_1: Not I; Site_2: Eco
                     RI; 1st strand cDNA was primed with a Not I - oligo(dT)
                     primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCGAATGGTTTTTTTTTTTTTTTTTTTT
                     T 3']; double-stranded cDNA was ligated to Eco RI
                     adaptors (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of the modified pT7T3 vector.
                     RNA provided by Dr. Minoru Ko, Wayne State Univ. Library
                     constructed and normalized by Bento Soares and M.Fatima
```

```
                    Bonaldo."
BASE COUNT       22 a     18 c     18 g     18 t
ORIGIN

  Query Match            71.0%;  Score 14.2;  DB 9;  Length 76;
  Best Local Similarity  84.2%;  Pred. No. 1.1e+04;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 CTGTGAGCTGGAAGTCATC 20
          ||||||||| || |||| |
Db      9 CTGTGAGCTTGAGGTCAGC 27


RESULT 9
BE957944
LOCUS       BE957944                 65 bp    mRNA    linear   EST 14-DEC-2000
DEFINITION  601654129R2 NIH_MGC_55 Homo sapiens cDNA clone IMAGE:3838950 3',
            mRNA sequence.
ACCESSION   BE957944
VERSION     BE957944.2  GI:11774721
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 65)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     On Oct 3, 2000 this sequence version replaced gi:10568649.
            Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: CLONETECH Laboratories, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM526  row: j  column: 07
            High quality sequence start: 29
            High quality sequence stop: 65.
FEATURES             Location/Qualifiers
     source          1. .65
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3838950"
                     /clone_lib="NIH_MGC_55"
                     /tissue_type="from acute myelogenous leukemia"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /note="Organ: bone marrow; Vector: pDNR-LIB (Clontech);
                     Site_1: SfiI (ggccgcctcggcc); Site_2: SfiI (ggccattatggcc
                     ); Double-stranded cDNA was prepared from cell line RNA.
                     5' and 3' adaptors were used in cloning as follows: 5'
                     adaptor sequence:  5'-CACGGCCATTATGGCC-3' and 3' adaptor
                     sequence:  5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)BN-3'
                     (where B = A, C, or G and N = A, C,  G, or T).   Average
                     insert size 1.65 kb (range 0.9-4.0 kb).  14/15 colonies
                     contained  inserts by PCR. This library was enriched for
                     full-length clones and was constructed by Clontech
                     Laboratories (Palo Alto, CA)."
BASE COUNT       13 a     11 c     22 g     19 t
ORIGIN

  Query Match            69.0%;  Score 13.8;  DB 12;  Length 65;
  Best Local Similarity  88.2%;  Pred. No. 1.6e+04;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      4 GTGAGCTGGAAGTCATC 20
          ||||||| ||| |||||
Db      1 GTGAGCTAGAACTCATC 17
```

```
RESULT 10
AZ467019/c
LOCUS       AZ467019                 68 bp    DNA     linear   GSS 04-OCT-2000
DEFINITION  1M0278E10F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0278E10 F, DNA sequence.
ACCESSION   AZ467019
VERSION     AZ467019.1  GI:10625144
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 68)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0278  row: E  column: 10
            Seq primer: CGTTGTAAAACGACGGCCAGT
            Class: plasmid ends
            High quality sequence stop: 68.
FEATURES             Location/Qualifiers
     source          1. .68
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC1M0278E10"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
BASE COUNT       24 a      8 c     14 g     22 t
ORIGIN

  Query Match            69.0%;  Score 13.8;  DB 17;  Length 68;
  Best Local Similarity  88.2%;  Pred. No. 1.7e+04;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 CCTGTGAGCTGGAAGTC 17
          ||||||| ||| |||||
Db     68 CCTGTGAACTGTAAGTC 52
```

```
RESULT 11
N67141
LOCUS       N67141                    79 bp    mRNA    linear   EST 18-MAR-1996
DEFINITION  yz46h01.s1 Morton Fetal Cochlea Homo sapiens cDNA clone
            IMAGE:286129 3' similar to gb:Z26876 60S RIBOSOMAL PROTEIN L38
            (HUMAN);, mRNA sequence.
ACCESSION   N67141
VERSION     N67141.1  GI:1219266
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 79)
  AUTHORS   Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiapelli,B.,
            Chissoe,S., Dietrich,N., DuBuque,T., Favello,A., Gish,W., Hawkins
            ,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N., Mardis,E., Moore
            ,B., Morris,M., Parsons,J., Prange,C., Rifkin,L., Rohlfing,T.,
            Schellenberg,K., Soares,M.B., Tan,F., Thierry-Meg,J., Trevaskis,E.,
            Underwood,K., Wohldmann,P., Waterston,R., Wilson,R. and Marra,M.
  TITLE     Generation and analysis of 280,000 human expressed sequence tags
  JOURNAL   Genome Res. 6 (9), 807-828 (1996)
  MEDLINE   97044478
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Trace considered overall poor quality
            Seq primer: m13 -40 forward
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .79
                     /organism="Homo sapiens"
                     /db_xref="GDB:3893246"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:286129"
                     /clone_lib="Morton Fetal Cochlea"
                     /tissue_type="cochlea"
                     /dev_stage="16-22 week fetus"
                     /lab_host="SOLR cells (kanamycin resistant)"
                     /note="Organ: ear; Vector: pBluescript SK-; Site_1: EcoRI;
                     Site_2: XhoI; Reference: Genomics 23, 42-50 (1994) Cloned
                     unidirectionally. Primer: Oligo dT. Fetal cochlea, normal.
                     37% of inserts <0.5 kb, 56% 0.5-1.0 kb, 7% >1 kb. Uni-ZAP
                     XR Vector. Library constructed by N. Robertson, C. Morton.
                     -5' adaptor sequence: 5' GAATTCGGCACGAG 3' -3' adaptor
                     sequence: 5' CTCGAGTTTTTTTTTTTTTTTTT 3'"
BASE COUNT        9 a     21 c     24 g     21 t      4 others
ORIGIN

  Query Match             69.0%;  Score 13.8;  DB 14;  Length 79;
  Best Local Similarity   83.3%;  Pred. No. 1.8e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 CTGTGAGCTGGAAGTCAT 19
          |||||||| ||||| | |
Db     21 CTGTGAGCCGGAAGNCCT 38


RESULT 12
BQ756524/c
LOCUS       BQ756524                  95 bp    mRNA    linear   EST 26-JUL-2002
DEFINITION  EBem09_SQ002_M05_R embryo, 1 Day germination, no treatment, cv
            Optic, EBem09 Hordeum vulgare cDNA clone EBem09_SQ002_M05 5', mRNA
            sequence.
ACCESSION   BQ756524
VERSION     BQ756524.1  GI:21964996
KEYWORDS    EST.
SOURCE      Hordeum vulgare.
  ORGANISM  Hordeum vulgare
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
            ; Triticeae; Hordeum.
REFERENCE   1  (bases 1 to 95)
  AUTHORS   Hedley,P., Liu,H., Caldwell,D., McCallum,N., Mudie,S., Cardle,L.,
            Ramsay,L., Machray,G., Marshall,D.F.M. and Waugh,R.
  TITLE     Development of Barley Transcriptome Resources
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Waugh R, Marshall DF
            Genome Dynamics/Computational Biology
            Scottish Crop Research Institute
            Invergowrie, Dundee, DD2 5DA, Scotland, UK
            Tel: 00 44 1382 562731
            Fax: 00 44 1382 562426
            Email: est@scri.sari.ac.uk.
FEATURES             Location/Qualifiers
     source          1. .95
                     /organism="Hordeum vulgare"
                     /cultivar="Optic"
                     /db_xref="taxon:4513"
                     /clone="EBem09_SQ002_M05"
                     /clone_lib="embryo, 1 Day germination, no treatment, cv
                     Optic, EBem09"
                     /tissue_type="embryo"
                     /dev_stage="1 Day germination"
                     /lab_host="DH10B"
                     /note="Vector: pSPORT1; Site_1: Sal I; Site_2: Not I;
                     Non-normalised library, directionally cloned into pSPORT1.
                     Derived from embryos dissected from germinating grains (1
                     day) in glasshouse grown barley plants. Developed as part
                     of the barley transcriptome resources of BBSRC/SEERAD
                     funded cereal IGF (Investigating Gene Function) project."
BASE COUNT       21 a     23 c     33 g     18 t
ORIGIN

  Query Match             69.0%;  Score 13.8;  DB 14;  Length 95;
  Best Local Similarity   88.2%;  Pred. No. 1.9e+04;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      4 GTGAGCTGGAAGTCATC 20
          | |||||||||| ||||
Db     48 GCGAGCTGGAAGGCATC 32


RESULT 13
R72481
LOCUS       R72481                    67 bp    mRNA    linear   EST 02-JUN-1995
DEFINITION  yj90f05.r1 Soares breast 2NbHBst Homo sapiens cDNA clone
            IMAGE:156033 5' similar to SP:SERR_DROME P18168 SERRATE PROTEIN
            PRECURSOR ;, mRNA sequence.
ACCESSION   R72481
VERSION     R72481.1  GI:846513
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 67)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman
            ,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J.,
            Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston
            ,R., Williamson,A., Wohldmann,P. and Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Insert Size: 924
            High quality sequence starts: 1
```

```
            High quality sequence stops: 1
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Trace considered overall poor quality
            Possible reversed clone: similarity on wrong strand
            Insert Length: 924   Std Error: 0.00
            Seq primer: M13RP1
            High quality sequence stop: 1.
FEATURES              Location/Qualifiers
     source           1. .67
                      /organism="Homo sapiens"
                      /db_xref="GDB:569675"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:156033"
                      /clone_lib="Soares breast 2NbHBst"
                      /sex="Female"
                      /dev_stage="adult"
                      /lab_host="DH10B (ampicillin resistant)"
                      /note="Organ: breast; Vector: pT7T3D (Pharmacia) with a
                      modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                      strand cDNA was primed with a Not I - oligo(dT) primer [5'
                      TGTTACCAATCTGAAGTGGGAGCGGCCGCCCTTTTTTTTTTTTTTTTTT 3'],
                      double-stranded cDNA was ligated to Eco RI adaptors
                      (Pharmacia), digested with Not I and cloned into the Not I
                      and Eco RI sites of a modified pT7T3 vector (Pharmacia).
                      Library went through one round of normalization to a Cot =
                      230. Library constructed by Bento Soares and M.Fatima
                      Bonaldo."
BASE COUNT       10 a      20 c      25 g      11 t       1 others
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 14;  Length 67;
  Best Local Similarity   80.0%;  Pred. No. 2.1e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 CCTGTGAGCTGGAAGTCATC 20
           ||||  ||||||| |||| |
Db      36 CCTGCCAGCTGGACGTCAAC 55


RESULT 14
BE318017/c
LOCUS       BE318017                  52 bp    mRNA    linear   EST 21-DEC-2000
DEFINITION  NF061E07LF1F1051 Developing leaf Medicago truncatula cDNA clone
            NF061E07LF 5', mRNA sequence.
ACCESSION   BE318017
VERSION     BE318017.2  GI:11961649
KEYWORDS    EST.
SOURCE      barrel medic.
  ORGANISM  Medicago truncatula
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
            Medicago.
REFERENCE   1  (bases 1 to 52)
  AUTHORS   Torres-Jerez,I., Scott,A.D., Harris,A.R., Gonzales,R.A., Bell,C.J.,
            Flores,H.R., Inman,J.T., Weller,J.W. and May,G.D.
  TITLE     Expressed Sequence Tags from the Samuel Roberts Noble Foundation
            Medicago truncatula leaf library
  JOURNAL   Unpublished (2000)
COMMENT     On Jul 14, 2000 this sequence version replaced gi:9191794.
            Contact: May GD
            Plant Biology Division
            The Samuel Roberts Noble Foundation
            2510 Sam Noble Parkway, Ardmore, OK 73402, USA
            Tel: 580 221 7391
            Fax: 580 221 7380
            Email: gdmay@noble.org
            Medicago Genome Initiative accession: MGI:S:22128
            Insert Length: 916   Std Error: 0.00
            Plate: 061  row: E  column: 07
            Seq primer: TCACACAGGAAACAGCTATGAC.
FEATURES              Location/Qualifiers
     source           1. .52
                      /organism="Medicago truncatula"
                      /db_xref="taxon:3880"
                      /clone="NF061E07LF"
                      /clone_lib="Developing leaf"
                      /tissue_type="leaf"
                      /dev_stage="Pooled developmental"
                      /note="Vector: Lambda Zap; Contains a mixture of very
                      young, developing, mature and senescing leaves."
BASE COUNT       10 a      21 c       8 g      13 t
ORIGIN

  Query Match             67.0%;  Score 13.4;  DB 10;  Length 52;
  Best Local Similarity   93.3%;  Pred. No. 2.3e+04;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 CCTGTGAGCTGGAAG 15
           ||||||| ||||||
Db      23 CCTGTGAGGTGGAAG 9


RESULT 15
D38680
LOCUS       D38680                    76 bp    mRNA    linear   EST 31-MAY-1995
DEFINITION  HUMCI215 Human chromosome 8 Homo sapiens cDNA 5', mRNA sequence.
ACCESSION   D38680
VERSION     D38680.1  GI:807739
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 76)
  AUTHORS   Koyama,K., Sudo,K. and Nakamura,Y.
  TITLE     Isolation of 115 human chromosome 8-specific expressed-sequence
            tags by exon amplification
  JOURNAL   Genomics 26, 245-253 (1995)
  MEDLINE   95324915
COMMENT     Contact: Yusuke Nakamura
            Institute of Medical Science
            University of Tokyo
            4-6-1, Shirokanedai, Minato-ku, Tokyo 108,  Japan
            Tel: 81-3-5449-5372
            Fax: 81-3-5449-5433
            Email: yusuke@ims.u-tokyo.ac.jp.
FEATURES              Location/Qualifiers
     source           1. .76
                      /organism="Homo sapiens"
                      /db_xref="taxon:9606"
                      /map="8"
                      /clone_lib="Human chromosome 8"
                      /note="Exon-like sequence on chromosome 8; exon-trapping
                      method"
BASE COUNT       19 a      15 c      26 g      16 t
ORIGIN

  Query Match             67.0%;  Score 13.4;  DB 14;  Length 76;
  Best Local Similarity   93.3%;  Pred. No. 2.7e+04;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       6 GAGCTGGAAGTCATC 20
           |||||||||| ||||
Db       1 GAGCTGGAAGGCATC 15


Search completed: November 26, 2002, 17:58:02
Job time : 808.5 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 3, 2002, 18:14:22 ; Search time 351.3 Seconds
(without alignments)
1656.863 Million cell updates/sec

Title: US-09-296-264-27
Perfect score: 20
Sequence: 1 catgtgataccagaaggtca 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 995600

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
c 1	18.6	93.0	21	6	AX097258	AX097258 Sequence
2	14.8	74.0	20	6	AX298848	AX298848 Sequence
c 3	14.2	71.0	51	5	CHKMYHF2	M18138 Chicken myo
4	13.6	68.0	53	6	AR062116	AR062116 Sequence
5	13.6	68.0	95	4	SSCASR3	AF041027 Sus scrof
6	13.2	66.0	54	6	AR080277	AR080277 Sequence
7	12.8	64.0	38	6	A08084	A08084 Synthetic o
c 8	12.8	64.0	43	6	A08081	A08081 Synthetic o
c 9	12.8	64.0	51	6	AX162992	AX162992 Sequence
c 10	12.6	63.0	24	6	AR037538	AR037538 Sequence
c 11	12.6	63.0	24	6	AX291677	AX291677 Sequence
12	12.6	63.0	26	6	A83974	A83974 Sequence 1
c 13	12.6	63.0	51	6	AX162994	AX162994 Sequence
c 14	12.6	63.0	72	6	A75006	A75006 Sequence 69
c 15	12.6	63.0	72	6	A77985	A77985 Sequence 69
c 16	12.6	63.0	84	6	A59475	A59475 Sequence 25
c 17	12.6	63.0	84	6	A59508	A59508 Sequence 58
18	12.6	63.0	86	6	AR140821	AR140821 Sequence
19	12.6	63.0	86	6	AR150771	AR150771 Sequence
20	12.6	63.0	86	6	I65649	I65649 Sequence 9
21	12.6	63.0	86	6	I67881	I67881 Sequence 9
22	12.6	63.0	86	6	I90102	I90102 Sequence 9
c 23	12.6	63.0	91	9	HUMMHDQB01	M27656 Homo sapien
c 24	12.6	63.0	96	9	AF307154	AF307154 Homo sapi
c 25	12.6	63.0	96	9	AF307156	AF307156 Homo sapi
c 26	12.6	63.0	96	9	AF506754	AF506754 Homo sapi
c 27	12.6	63.0	96	9	AY036896	AY036896 Homo sapi
c 28	12.6	63.0	96	9	AY036897	AY036897 Homo sapi
29	12.4	62.0	20	6	AR016217	AR016217 Sequence
30	12.4	62.0	20	6	AR035577	AR035577 Sequence
c 31	12.4	62.0	20	6	AR035582	AR035582 Sequence
32	12.4	62.0	20	6	AR035642	AR035642 Sequence
33	12.4	62.0	20	6	I84436	I84436 Sequence 4
c 34	12.4	62.0	20	6	I84441	I84441 Sequence 9
c 35	12.4	62.0	75	6	AR145217	AR145217 Sequence
c 36	12.4	62.0	90	6	AX384698	AX384698 Sequence
37	12.4	62.0	90	6	AX384704	AX384704 Sequence
c 38	12.2	61.0	25	6	AR143580	AR143580 Sequence
c 39	12.2	61.0	25	6	AR168949	AR168949 Sequence
40	12.2	61.0	35	6	A42707	A42707 Sequence 39
41	12.2	61.0	35	6	A42708	A42708 Sequence 40
42	12.2	61.0	35	6	I87213	I87213 Sequence 39
43	12.2	61.0	35	6	I87214	I87214 Sequence 40
c 44	12.2	61.0	36	6	AR067852	AR067852 Sequence
c 45	12.2	61.0	36	6	AR106653	AR106653 Sequence

ALIGNMENTS

RESULT 1
AX097258/c
LOCUS AX097258 21 bp DNA linear PAT 30-MAR-2001
DEFINITION Sequence 2436 from Patent WO0118250.
ACCESSION AX097258
VERSION AX097258.1 GI:13513668
KEYWORDS .
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 21)
AUTHORS Lander,E.S., Gargill,M., Ireland,J.S., Bolk,S., Daley,G.Q. and
Mccarthy,J.J.
TITLE Single nucleotide polymorphisms in genes

```
  JOURNAL   Patent: WO 0118250-A 2436 15-MAR-2001;
            WHITEHEAD INSTITUTE FOR BIOMEDICAL RESEARCH (US) ; Millennium
            Pharmaceuticals, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .21
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT        4 a      6 c      4 g      6 t      1 others
ORIGIN

  Query Match             93.0%;  Score 18.6;  DB 6;  Length 21;
  Best Local Similarity   94.7%;  Pred. No. 80;
  Matches   18;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 CATGTGATACCAGAAGGTC 19
          |||||||||:||||||||||
Db     19 CATGTGATRCCAGAAGGTC 1


RESULT 2
AX298848
LOCUS       AX298848                  20 bp    DNA     linear   PAT 26-NOV-2001
DEFINITION  Sequence 482 from Patent WO0183749.
ACCESSION   AX298848
VERSION     AX298848.1  GI:17128838
KEYWORDS    .
SOURCE      Mus sp.
  ORGANISM  Mus sp.
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Bachmanov,A.A., Beauchamp,G.K., Chatterjee,A., de Jong,P.J., Li,S.,
            Li,X., Ohmen,J.D., Reed,D.R., Ross,D. and Tordoff,M.G.
  TITLE     Gene and sequence variation associated with sensing carbohydrate
            compounds and other sweeteners
  JOURNAL   Patent: WO 0183749-A 482 08-NOV-2001;
            WARNER-LAMBERT COMPANY (US) ; The Monell Chemical Senses Center
            (US)
FEATURES             Location/Qualifiers
     source          1. .20
                     /organism="Mus sp."
                     /db_xref="taxon:10095"
BASE COUNT        7 a      4 c      7 g      2 t
ORIGIN

  Query Match             74.0%;  Score 14.8;  DB 6;  Length 20;
  Best Local Similarity   88.9%;  Pred. No. 8.8e+03;
  Matches   16;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 CATGTGATACCAGAAGGT 18
          |||| || ||||||||||
Db      2 CATGGGAGACCAGAAGGT 19


RESULT 3
CHKMYHF2/c
LOCUS       CHKMYHF2                  51 bp    DNA     linear   VRT 28-APR-1993
DEFINITION  Chicken myosin heavy chain gene, exon 2.
ACCESSION   M18138 J03467
VERSION     M18138.1  GI:212378
KEYWORDS    myosin; myosin heavy chain.
SEGMENT     2 of 3
SOURCE      Chicken DNA, clone N127.
  ORGANISM  Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE   1  (bases 1 to 51)
  AUTHORS   Kropp,K.E., Gulick,J. and Robbins,J.
  TITLE     Structural and transcriptional analysis of a chicken myosin heavy
            chain gene subset
  JOURNAL   J. Biol. Chem. 262 (34), 16536-16545 (1987)
  MEDLINE   88059036
    PUBMED  2824498
FEATURES             Location/Qualifiers
     source          1. .51
                     /organism="Gallus gallus"
                     /db_xref="taxon:9031"
     prim_transcript <1. .>51
                     /note="myosin heavy chain mRNA and introns"
     intron          <1. .2
                     /note="intron A"
     intron          50. .>51
                     /note="intron B"
BASE COUNT       13 a     14 c     10 g     14 t
ORIGIN      161 bp after segment 1.

  Query Match             71.0%;  Score 14.2;  DB 5;  Length 51;
  Best Local Similarity   84.2%;  Pred. No. 1.7e+04;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 CATGTGATACCAGAAGGTC 19
          | ||||| || ||||||||
Db     47 CTTGTGACACAAGAAGGTC 29


RESULT 4
AR062116
LOCUS       AR062116                  53 bp    DNA     linear   PAT 29-SEP-1999
DEFINITION  Sequence 3 from patent US 5843678.
ACCESSION   AR062116
VERSION     AR062116.1  GI:5989807
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 53)
  AUTHORS   Boyle,W.J.
  TITLE     Osteoprotegerin binding proteins
  JOURNAL   Patent: US 5843678-A 3 01-DEC-1998;
FEATURES             Location/Qualifiers
     source          1. .53
                     /organism="unknown"
BASE COUNT       16 a     18 c      5 g     14 t
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 6;  Length 53;
  Best Local Similarity   80.0%;  Pred. No. 3.5e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 CATGTGATACCAGAAGGTCA 20
          ||| ||| ||| |||| |||
Db      9 CATATGAAACCTGAAGCTCA 28


RESULT 5
SSCASR3
LOCUS       SSCASR3                   95 bp    DNA     linear   MAM 03-SEP-1999
DEFINITION  Sus scrofa calcium-sensing receptor (CASR) gene, exon 7 and partial
            cds.
ACCESSION   AF041027
VERSION     AF041027.1  GI:3098469
KEYWORDS    .
SEGMENT     3 of 3
SOURCE      Sus scrofa.
  ORGANISM  Sus scrofa
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE   1  (bases 1 to 95)
  AUTHORS   Van Poucke,M., Tornsten,A., Mattheeuws,M., Van Zeveren,A.,
            Peelman,L.J. and Chowdhary,B.P.
  TITLE     Comparative mapping between human chromosome 3 and porcine
            chromosome 13
  JOURNAL   Cytogenet. Cell Genet. 85 (3-4), 279-284 (1999)
  MEDLINE   99380101
   PUBMED   10449918
```

```
REFERENCE   2  (bases 1 to 95)
  AUTHORS   Van Poucke,M. and Peelman,L.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-JAN-1998) Animal Genetics, University of Ghent,
            Heidestraat, Merelbeke 9820, Belgium
FEATURES             Location/Qualifiers
     source          1. .95
                     /organism="Sus scrofa"
                     /strain="Belgian Landrace"
                     /db_xref="taxon:9823"
                     /chromosome="13"
                     /map="13q41"
     gene            order(AF041025.1:<1. .355,AF041026.1:1. .178,1. .>95)
                     /gene="CASR"
     mRNA            order(AF041025.1:<1. .355,AF041026.1:1. .178,1. .>95)
                     /gene="CASR"
                     /product="calcium-sensing receptor"
     CDS             <1. .>95
                     /gene="CASR"
                     /codon_start=2
                     /product="calcium-sensing receptor"
                     /protein_id="AAC15662.1"
                     /db_xref="GI:3098473"
                     /translation="QQQPRCKQKVIFGSGTVTFSLSFHEPQKNAM"
     exon            <1. .>95
                     /gene="CASR"
                     /number=7
BASE COUNT       24 a     26 c     26 g     19 t
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 4;  Length 95;
  Best Local Similarity   80.0%;  Pred. No. 3.3e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 CATGTGATACCAGAAGGTCA 20
          || |||  | ||||||||||
Db     13 CAGGTGTAAGCAGAAGGTCA 32


RESULT 6
AR080277
LOCUS       AR080277                  54 bp    DNA     linear   PAT 31-AUG-2000
DEFINITION  Sequence 11 from patent US 5968750.
ACCESSION   AR080277
VERSION     AR080277.1  GI:10007012
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 54)
  AUTHORS   Zolotukhin,S., Muzyczka,N. and Hauswirth,W.W.
  TITLE     Humanized green fluorescent protein genes and methods
  JOURNAL   Patent: US 5968750-A 11 19-OCT-1999;
FEATURES             Location/Qualifiers
     source          1. .54
                     /organism="unknown"
BASE COUNT       15 a     12 c     15 g     12 t
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 6;  Length 54;
  Best Local Similarity   83.3%;  Pred. No. 5.7e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 CATGTGATACCAGAAGGT 18
          || |  ||||||||||||
Db     35 CACGCCATACCAGAAGGT 52


RESULT 7
A08084
LOCUS       A08084                    38 bp    DNA     linear   PAT 20-JUL-1993
DEFINITION  Synthetic oligonucleotide D13a.
ACCESSION   A08084
VERSION     A08084.1  GI:413326
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 38)
  AUTHORS   Heinzel-Wieland,R., Ammann,J., Steffens,G.J. and Flohe,L.
  TITLE     Serine protease inhibitor proteins, medicaments containing them,
            DNA sequences coding for these proteins and methods for producing
            these proteins, medicaments and DNA sequences
  JOURNAL   Patent: EP 0373335-A 6 20-JUN-1990;
            Gruenenthal GmbH
FEATURES             Location/Qualifiers
     source          1. .38
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
BASE COUNT       14 a      9 c      8 g      7 t
ORIGIN

  Query Match             64.0%;  Score 12.8;  DB 6;  Length 38;
  Best Local Similarity   87.5%;  Pred. No. 9.7e+04;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 CATGTGATACCAGAAG 16
          ||| |||||||| |||
Db      9 CATTTGATACCACAAG 24


RESULT 8
A08081/c
LOCUS       A08081                    43 bp    DNA     linear   PAT 20-JUL-1993
DEFINITION  Synthetic oligonucleotide D13.
ACCESSION   A08081
VERSION     A08081.1  GI:413323
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 43)
  AUTHORS   Heinzel-Wieland,R., Ammann,J., Steffens,G.J. and Flohe,L.
  TITLE     Serine protease inhibitor proteins, medicaments containing them,
            DNA sequences coding for these proteins and methods for producing
            these proteins, medicaments and DNA sequences
  JOURNAL   Patent: EP 0373335-A 3 20-JUN-1990;
            Gruenenthal GmbH
FEATURES             Location/Qualifiers
     source          1. .43
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
BASE COUNT       10 a      8 c     11 g     14 t
ORIGIN

  Query Match             64.0%;  Score 12.8;  DB 6;  Length 43;
  Best Local Similarity   87.5%;  Pred. No. 9.6e+04;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 CATGTGATACCAGAAG 16
          ||| |||||||| |||
Db     39 CATTTGATACCACAAG 24


RESULT 9
AX162992/c
LOCUS       AX162992                  51 bp    DNA     linear   PAT 22-JUN-2001
DEFINITION  Sequence 6320 from Patent WO0140521.
ACCESSION   AX162992
VERSION     AX162992.1  GI:14544323
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 51)
```

```
  AUTHORS   Shimkets,R.A. and Leach,M.
  TITLE     Nucleic acids containing single nucleotide polymorphisms and
            methods of use thereof
  JOURNAL   Patent: WO 0140521-A 6320 07-JUN-2001;
            Curagen Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .51
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     misc_feature    26
                     /note="2 of 2 allelic variants (6319 is other entry)
                     Accession number cg42716656"
BASE COUNT       13 a     11 c     11 g     16 t
ORIGIN

  Query Match             64.0%;  Score 12.8;  DB 6;  Length 51;
  Best Local Similarity   87.5%;  Pred. No. 9.4e+04;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       5 TGATACCAGAAGGTCA 20
            || ||| |||||||||
Db      33 TGTTACTAGAAGGTCA 18


RESULT 10
AR037538/c
LOCUS       AR037538                  24 bp    DNA     linear   PAT 29-SEP-1999
DEFINITION  Sequence 16 from patent US 5800982.
ACCESSION   AR037538
VERSION     AR037538.1  GI:5956255
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 24)
  AUTHORS   Hasegawa,A., Maki,N., Yagi,S., Kashiwakuma,T., Yamaguchi,K.,
            Ikeguchi,N., Kobayashi,T. and Senoo,C.
  TITLE     Antigenic peptides for growing hepatitis C virus, kit comprising
            the same and methods for its grouping using the same
  JOURNAL   Patent: US 5800982-A 16 01-SEP-1998;
FEATURES             Location/Qualifiers
     source          1. .24
                     /organism="unknown"
BASE COUNT        7 a      6 c      4 g      7 t
ORIGIN

  Query Match             63.0%;  Score 12.6;  DB 6;  Length 24;
  Best Local Similarity   78.9%;  Pred. No. 1.3e+05;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 CATGTGATACCAGAAGGTC 19
            |||||| || ||| || ||
Db      19 CATGTGTTAGCAGTAGATC 1


RESULT 11
AX291677/c
LOCUS       AX291677                  24 bp    DNA     linear   PAT 21-NOV-2001
DEFINITION  Sequence 3439 from Patent WO0179548.
ACCESSION   AX291677
VERSION     AX291677.1  GI:17053360
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Barany,F., Zirvi,M., Gerry,N.P., Favis,R. and Kliman,R.
  TITLE     Method of designing addressable array for detection of nucleic acid
            sequence differences using ligase detection reaction
  JOURNAL   Patent: WO 0179548-A 3439 25-OCT-2001;
            CORNELL RESEARCH FOUNDATION, INC. (US)
FEATURES             Location/Qualifiers
     source          1. .24
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="Hypothetical Probe Sequence"
BASE COUNT        3 a      7 c      8 g      6 t
ORIGIN

  Query Match             63.0%;  Score 12.6;  DB 6;  Length 24;
  Best Local Similarity   78.9%;  Pred. No. 1.3e+05;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 CATGTGATACCAGAAGGTC 19
            || | |||||||  |||||
Db      19 CACGGGATACCACGAGGTC 1


RESULT 12
A83974
LOCUS       A83974                    26 bp    DNA     linear   PAT 21-JAN-2000
DEFINITION  Sequence 1 from Patent WO9846792.
ACCESSION   A83974
VERSION     A83974.1  GI:6733114
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 26)
  AUTHORS   Chang,K.C.
  TITLE     GENETIC MARKER BASED PIG SELECTION
  JOURNAL   Patent: WO 9846792-A 1 22-OCT-1998;
            CHANG KIN CHOW (GB); BIOTECH & BIOLOG SCIEN RES (GB)
FEATURES             Location/Qualifiers
     source          1. .26
                     /organism="unidentified"
                     /db_xref="taxon:32644"
BASE COUNT        9 a      5 c      7 g      5 t
ORIGIN

  Query Match             63.0%;  Score 12.6;  DB 6;  Length 26;
  Best Local Similarity   78.9%;  Pred. No. 1.3e+05;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 CATGTGATACCAGAAGGTC 19
            ||||||| || |||| | |
Db       1 CATGTGAGACTAGAATGCC 19


RESULT 13
AX162994/c
LOCUS       AX162994                  51 bp    DNA     linear   PAT 22-JUN-2001
DEFINITION  Sequence 6322 from Patent WO0140521.
ACCESSION   AX162994
VERSION     AX162994.1  GI:14544325
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 51)
  AUTHORS   Shimkets,R.A. and Leach,M.
  TITLE     Nucleic acids containing single nucleotide polymorphisms and
            methods of use thereof
  JOURNAL   Patent: WO 0140521-A 6322 07-JUN-2001;
            Curagen Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .51
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     misc_feature    26
                     /note="2 of 2 allelic variants (6321 is other entry)
                     Accession number cg42716656"
BASE COUNT       16 a     16 c      8 g     11 t
ORIGIN
```

```
  Query Match             63.0%;  Score 12.6;  DB 6;  Length 51;
  Best Local Similarity   78.9%;  Pred. No. 1.2e+05;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      2 ATGTGATACCAGAAGGTCA 20
          || || |||  ||||||||
Db     27 ATCTGTTACTGGAAGGTCA 9


RESULT 14
A75006/c
LOCUS       A75006                    72 bp    DNA     linear   PAT 15-OCT-1999
DEFINITION  Sequence 692 from Patent WO9401548.
ACCESSION   A75006
VERSION     A75006.1  GI:6065019
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 72)
  AUTHORS   Sibson,D.R. and Gross,J.
  TITLE     HUMAN NUCLEIC ACID FRAGMENTS, ISOLATED FROM BRAIN ADRENAL TISSUE,
            PLACENTA OR BONE NARROW
  JOURNAL   Patent: WO 9401548-A 692 20-JAN-1994;
            MEDICAL RES COUNCIL (GB); SIBSON DAVID ROSS (GB)
FEATURES             Location/Qualifiers
     source          1. .72
                     /organism="unidentified"
                     /db_xref="taxon:32644"
BASE COUNT       27 a     17 c     17 g     11 t
ORIGIN

  Query Match             63.0%;  Score 12.6;  DB 6;  Length 72;
  Best Local Similarity   78.9%;  Pred. No. 1.2e+05;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 CATGTGATACCAGAAGGTC 19
          |||| | | ||||||||| |
Db     20 CATGGGTTTCCAGAAGGCC 2


RESULT 15
A77985/c
LOCUS       A77985                    72 bp    DNA     linear   PAT 19-OCT-1999
DEFINITION  Sequence 692 from Patent EP0587279.
ACCESSION   A77985
VERSION     A77985.1  GI:6089650
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 72)
  AUTHORS   Sibson,D.R. and Hadfield,K.M.
  TITLE     HUMAN NUCLEIC ACID FRAGMENTS ISOLATED FROM BRAIN, ADRENAL TISSUE,
            PLACENTA OR BONE NARROW AND THEIR USE
  JOURNAL   Patent: EP 0587279-A 692 16-MAR-1994;
            MEDICAL RES COUNCIL (GB)
FEATURES             Location/Qualifiers
     source          1. .72
                     /organism="unidentified"
                     /db_xref="taxon:32644"
BASE COUNT       27 a     17 c     17 g     11 t
ORIGIN

  Query Match             63.0%;  Score 12.6;  DB 6;  Length 72;
  Best Local Similarity   78.9%;  Pred. No. 1.2e+05;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 CATGTGATACCAGAAGGTC 19
          |||| | | ||||||||| |
Db     20 CATGGGTTTCCAGAAGGCC 2


Search completed: December  3, 2002, 22:23:38
Job time : 356.3 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 06:29:45 ; Search time 94.1 Seconds
(without alignments)
478.639 Million cell updates/sec

Title: US-09-296-264-27
Perfect score: 20
Sequence: 1 catgtgataccagaaggtca 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 2390332

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	20	100.0	20	21	AAZ31457	Human neuropilin m
c	2	19	95.0	21	22	AAF97671	Human gene single
	3	14.8	74.0	20	24	AAS97872	Murine SAC1 gene-s
c	4	13.8	69.0	27	16	AAQ81253	Ribozyme target se
c	5	13.8	69.0	51	22	AAL27376	Human SNP oligonuc
	6	13.6	68.0	21	12	AAQ12617	Sense primer DRS18
	7	13.6	68.0	41	24	AAS19393	Human gap joining
	8	13.6	68.0	53	19	AAV70296	Murine osteoproteg
c	9	13.6	68.0	60	24	ABN35867	Human spliced tran
c	10	13.4	67.0	23	24	ABK50854	Lentinula edodes p
c	11	13.4	67.0	60	24	ABN32733	Human spliced tran
	12	13.4	67.0	60	24	ABN40118	Human spliced tran
	13	13.2	66.0	54	20	AAX08460	Primer for amplify
	14	13.2	66.0	65	24	ABN54618	Mouse spliced tran
c	15	13.2	66.0	65	24	ABN54889	Mouse spliced tran
c	16	12.8	64.0	20	24	ABQ62407	Human syntaxin 4 i
c	17	12.8	64.0	51	22	AAI79379	Human silent SNP c
	18	12.8	64.0	60	24	ABN33745	Human spliced tran
c	19	12.8	64.0	60	24	ABN37892	Human spliced tran
	20	12.8	64.0	60	24	ABN47018	Human spliced tran
c	21	12.6	63.0	24	15	AAQ58469	HCV peptide C14-1-
c	22	12.6	63.0	24	24	ABI89274	Capture oligonucle
	23	12.6	63.0	24	24	ABI89275	Capture oligonucle
	24	12.6	63.0	26	19	AAV08293	PCR primer for por
c	25	12.6	63.0	36	15	AAQ74189	3' end fragment of
c	26	12.6	63.0	45	12	AAQ10551	Bone calcification
c	27	12.6	63.0	51	22	AAI79381	Human silent SNP c
	28	12.6	63.0	53	24	AAL40727	Human IRS-2 gene r
c	29	12.6	63.0	60	24	ABN40107	Human spliced tran
	30	12.6	63.0	60	24	ABN42910	Human spliced tran
	31	12.6	63.0	60	24	ABN48666	Human spliced tran
	32	12.6	63.0	60	24	ABN50139	Human spliced tran
	33	12.6	63.0	63	21	AAA61455	HBV amplification
	34	12.6	63.0	63	22	AAH41821	Hepatitis B virus
c	35	12.6	63.0	65	24	ABN31069	Rat spliced transc
	36	12.6	63.0	65	24	ABN52868	Mouse spliced tran
	37	12.6	63.0	65	24	ABN54016	Mouse spliced tran
	38	12.6	63.0	65	24	ABN54731	Mouse spliced tran
c	39	12.6	63.0	72	15	AAQ77092	Human genome fragm
	40	12.6	63.0	78	18	AAT50934	Mouse p53-recognit
c	41	12.6	63.0	84	18	AAT84373	Friedreich's ataxi
	42	12.6	63.0	86	18	AAT65225	Platelet derived g
	43	12.6	63.0	86	20	AAX87009	Platelet derived g
	44	12.4	62.0	20	18	AAT68382	Loci-specific prim
	45	12.4	62.0	20	19	AAV00646	Primer 3115 used t

ALIGNMENTS

RESULT 1
AAZ31457
ID AAZ31457 standard; DNA; 20 BP.
XX
AC AAZ31457;
XX
DT 07-FEB-2000 (first entry)
XX
DE Human neuropilin mRNA specific antisense oligo GTI3628.
XX
KW Neuropilin; human; growth; metastasis; tumor; neovascularisation;
KW cancer; papilloma; diabetic retinopathy; antisense; ss.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9955855-A2.
XX
PD 04-NOV-1999.
XX
PF 23-APR-1999; 99WO-CA00324.
XX
PR 23-APR-1998; 98US-0082791.
XX
PA (GENE-) GENESENSE TECHNOLOGIES INC.
XX
PI Wright JA, Young AH, Lee YS;
XX
DR WPI; 2000-023357/02.
XX
PT Antisense oligonucleotides that inhibit neuropilin expression, useful
PT for treating cancer -

```
XX
PS   Claim 4; Page 17; 57pp; English.
XX
CC   Sequences AAZ31431-460 represent antisense oligonucleotides which
CC   inhibit human neuropilin expression. The antisense oligonucleotides can
CC   be used to inhibit the growth or metastasis of a mammalian tumor and
CC   inhibit neovascularisation. The oligonucleotides may be used to treat
CC   various forms of cancers or tumors, such as sarcomas, melanomas,
CC   adenomas, carcinomas of solid tissue, hypoxic tumors, squamous cell
CC   carcinomas of the mouth, throat, larynx and lung, genitourinary cancers
CC   such as cervical and bladder cancer, hematopoietic cancers, colon cancer,
CC   breast cancer, pancreatic cancer, renal cancer, brain cancer, skin
CC   cancer, liver cancer, head and neck cancers, and nervous system cancers,
CC   as well as benign lesions such as papillomas. The methods may be used to
CC   treat neovascularisation disorders such as diabetic retinopathy, and
CC   retinopathy of prematurity and age related macular degeneration.
XX
SQ   Sequence 20 BP; 7 A; 4 C; 5 G; 4 T; 0 other;

  Query Match             100.0%;  Score 20;  DB 21;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 0.49;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 CATGTGATACCAGAAGGTCA 20
          ||||||||||||||||||||
Db      1 CATGTGATACCAGAAGGTCA 20


RESULT 2
AAF97671/c
ID   AAF97671 standard; DNA; 21 BP.
XX
AC   AAF97671;
XX
DT   06-JUN-2001  (first entry)
XX
DE   Human gene single nucleotide polymorphism #2432.
XX
KW   Human; variant thrombospondin 1; variant thrombospondin 4; SNP;
KW   polymorphism; vascular disease; coronary artery disease; forensics;
KW   myocardial infarction; atherosclerosis; stroke; venous thromboembolism;
KW   pulmonary embolism; paternity test; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Variation       replace(11,C)
FT                   /*tag=  a
FT                   /standard_name= "single nucleotide polymorphism"
XX
PN   WO200118250-A2.
XX
PD   15-MAR-2001.
XX
PF   07-SEP-2000; 2000WO-US24503.
XX
PR   10-SEP-1999;   99US-0153357.
PR   26-JUL-2000; 2000US-0220947.
PR   16-AUG-2000; 2000US-0225724.
XX
PA   (WHED ) WHITEHEAD INST BIOMEDICAL RES.
PA   (MILL-) MILLENNIUM PHARM INC.
XX
PI   Lander ES,  Gargill M,  Ireland JS,  Bolk S,  Daley GQ,  McCarthy JJ;
XX
DR   WPI; 2001-226749/23.
XX
PT   Nucleic acids comprising single nucleotide polymorphisms, useful in
PT   applications such as forensics, paternity testing, medicine, genetic
PT   analysis and phenotype correlations to diseases such as diabetes and
PT   atherosclerosis -
XX
PS   Examples; Page 213; 242pp; English.
XX
CC   The present invention provides a method of diagnosing a vascular disease
CC   in an individual, involving determining the sequence at various
CC   polymorphic sites within the human thrombospondin 1 and thrombospondin 4
CC   genes. The sequences at a number of polymorphic sites are also provided
CC   in the specification. In particular, the method can be used in the
CC   diagnosis of atherosclerosis, myocardial infarction, coronary heart
CC   disease, stroke, peripheral vascular diseases, venous thromboembolism
CC   and pulmonary embolism. Single nucleotide polymorphisms (SNPs) are also
CC   useful in forensics, paternity testing, genetic analysis and phenotype
CC   correlations to diseases. The present sequence is an example of one of
CC   the human gene SNPS shown in the specification.
XX
SQ   Sequence 21 BP; 4 A; 6 C; 4 G; 7 T; 0 other;

  Query Match             95.0%;  Score 19;  DB 22;  Length 21;
  Best Local Similarity   100.0%;  Pred. No. 1.7;
  Matches   19;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 CATGTGATACCAGAAGGTC 19
          |||||||||||||||||||
Db     19 CATGTGATACCAGAAGGTC 1


RESULT 3
AAS97872
ID   AAS97872 standard; DNA; 20 BP.
XX
AC   AAS97872;
XX
DT   12-MAR-2002  (first entry)
XX
DE   Murine SAC1 gene-specific oligonucleotide PCR primer #439.
XX
KW   Human; mouse; SAC1; carbohydrate; sweetener; ethanol; alcoholism; ss;
KW   obesity; diabetes; transgenic embryo; body tissue; body fluid; pancreas;
KW   blood; tongue; PCR primer; anorectic; antidiabetic; gene therapy;
KW   protein replacement therapy.
XX
OS   Mus sp.
XX
PN   WO200183749-A2.
XX
PD   08-NOV-2001.
XX
PF   25-APR-2001; 2001WO-US13387.
XX
PR   28-APR-2000; 2000US-200794P.
PR   28-JUL-2000; 2000US-221419P.
PR   10-NOV-2000; 2000US-247443P.
XX
PA   (WARN ) WARNER LAMBERT CO.
PA   (MONE-) MONELL CHEM SENSES CENT.
XX
PI   Bachmanov AA,  Beauchamp GK,  Chatterjee A,  De Jong PJ,  Li S,  Li X;
PI   Ohmen JD,  Reed DR,  Ross D,  Tordoff MG;
XX
DR   WPI; 2002-075162/10.
XX
PT   Novel isolated polypeptide comprising variant form of mouse or human
PT   SAC1 polypeptide, and is associated with altered preference for
PT   carbohydrates or other sweeteners, useful for preventing obesity,
PT   diabetes, alcoholism  -
XX
PS   Claim 14; Page 91; 239pp; English.
XX
CC   The invention relates to an isolated polypeptide, comprising a variant
CC   form of mouse or human SAC1 polypeptide. The variant form is associated
CC   with altered preference for carbohydrates, other sweeteners or ethanol.
CC   The polypeptide and its associated DNA sequence can be produced by
CC   recombinant techniques and is useful for preventing obesity, diabetes or
CC   alcoholism associated with SAC1 expression. The sequences are useful in
CC   screening for drugs and sweeteners. Recombinant cell lines and transgenic
```

```
CC   embryos may be used in screening for and identifying agents that induce
CC   or repress function of SAC1. Predisposition to diabetes, obesity or
CC   alcoholism can be ascertained by testing any fluid or tissue of a human
CC   (such as blood, pancreas or tongue) for sequence variations of the SAC1
CC   gene. A sequence variation of the SAC1 locus may indicate a
CC   predisposition to diabetes, obesity and/or alcoholism and may provide a
CC   diagnostic mark. The polynucleotide can be detected in a biological
CC   sample by contacting the DNA with a probe to form a hybridisation complex
CC   which is then detected. The sequences represent cDNA encoding human and
CC   mouse SAC1 polypeptides and PCR primers specific for the SCA1 genes.
XX
SQ   Sequence 20 BP; 7 A; 4 C; 7 G; 2 T; 0 other;

  Query Match             74.0%;  Score 14.8;  DB 24;  Length 20;
  Best Local Similarity   88.9%;  Pred. No. 2.9e+02;
  Matches   16;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 CATGTGATACCAGAAGGT 18
           |||| || ||||||||||
Db       2 CATGGGAGACCAGAAGGT 19


RESULT 4
AAQ81253/c
ID   AAQ81253 standard; mRNA; 27 BP.
XX
AC   AAQ81253;
XX
DT   07-SEP-1995  (first entry)
XX
DE   Ribozyme target sequence in TGF-beta mRNA (bases 608-634).
XX
KW   Target site; ribozyme; hammerhead; hairpin; hepatitis delta virus;
KW   group 1 intron; RNaseP RNA motif; transforming growth factor-beta;
KW   TGF-beta; fibrous; connective; tissue disease; TGF-alpha; inhibin;
KW   epidermal growth factor; EGF; activin; amphiregulin; insulin;
KW   bone morphogenic protein; fibroblast growth factor; relaxin; ss.
XX
OS   Homo sapiens.
XX
PN   WO9429452-A.
XX
PD   22-DEC-1994.
XX
PF   02-JUN-1994;   94WO-US06331.
XX
PR   09-JUN-1993;   93US-0074343.
XX
PA   (RIBO-) RIBOZYME PHARM INC.
XX
PI   Draper KG;
XX
DR   WPI; 1995-051612/07.
XX
PT   Enzymatic RNA molecule with, e.g. a hammerhead or hairpin motif
PT   - cleaves mRNA associated with fibrous or connective tissue
PT   disease, and is useful for treatment or prophylaxis of such
PT   diseases
XX
PS   Claim 3; Page 4; 63pp; English.
XX
CC   The sequences (AAQ81238-304) represent the target sites where a ribozyme
CC   (hammerhead, hairpin, hepatitis delta virus, group 1 intron or RNaseP
CC   RNA motif) cleaves the mRNA of the transforming growth factor-beta
CC   (TGF-beta) gene.  This sequence corresponds to bases 608-634 of the
CC   TGF-beta mRNA. The ribozymes can also target the mRNAs of genes
CC   associated with the development or maintenance of fibrous or connective
CC   tissue disease in order to prevent or treat these diseases.  Such genes
CC   include TGF-alpha or beta, epidermal growth factor, inhibins, activins,
CC   amphiregulin, bone morphogenic proteins, fibroblast growth factors a and
CC   b, insulin growth factor 1 or 2, insulin or relaxin.
XX
SQ   Sequence 27 BP; 4 A; 13 C; 5 G; 5 U; 0 other;

  Query Match             69.0%;  Score 13.8;  DB 16;  Length 27;
  Best Local Similarity   88.2%;  Pred. No. 1e+03;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       2 ATGTGATACCAGAAGGT 18
           || || ||||||||||
Db      20 ATCTGGTACCAGAAGGT 4


RESULT 5
AAL27376/c
ID   AAL27376 standard; DNA; 51 BP.
XX
AC   AAL27376;
XX
DT   24-JAN-2002  (first entry)
XX
DE   Human SNP oligonucleotide #584.
XX
KW   Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
KW   neuroprotective; antimicrobial; gene therapy; vaccine; amylase; cancer;
KW   amyloid protein; angiopoietin; apoptosis related protein; cadherin;
KW   cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
KW   complement related protein; cytochrome; kinesin; cytokine; interferon;
KW   interleukin; G-protein coupled receptor; thioesterase; inflammation;
KW   multifactorial disease; autoimmune disease; infection;
KW   nervous system disease; ss.
XX
OS   Homo sapiens.
XX
PN   WO200147944-A2.
XX
PD   05-JUL-2001.
XX
PF   28-DEC-2000; 2000WO-US35498.
XX
PR   28-DEC-1999;   99US-0173419.
PR   27-DEC-2000; 2000US-0173419.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Shimkets RA,  Leach M;
XX
DR   WPI; 2001-465210/50.
XX
PT   Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
PT   oncogenes and histones, useful for diagnosing and treating, e.g.
PT   cancer, autoimmune diseases and infections -
XX
PS   Claim 1; Page 1555; 4143pp; English.
XX
CC   The present invention relates to oligonucleotides encoding polymorphic
CC   variants of proteins related to amylases, amyloid proteins, angiopoietin,
CC   apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,
CC   histones, kinases, colony stimulating factors, complement related
CC   proteins, cytochromes, kinesins, cytokines, interferons, interleukins,
CC   G-protein coupled receptors and thioesterases. The present sequence is
CC   one such oligonucleotide. The oligonucleotides and the peptides encoded
CC   by them may be used in the prevention, diagnosis and treatment of
CC   diseases associated with inappropriate expression of the proteins listed
CC   above. Disorders that may be prevented, diagnosed and/or treated include
CC   multifactorial diseases with a genetic component, such as autoimmune
CC   diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,
CC   systemic lupus erythromatosus and Grave's disease), inflammation, cancer
CC   (e.g. cancers of the bladder, brain, breast, colon and kidney,
CC   leukaemia), diseases of the nervous system and an infection of pathogenic
CC   organisms.
XX
SQ   Sequence 51 BP; 6 A; 19 C; 13 G; 13 T; 0 other;

  Query Match             69.0%;  Score 13.8;  DB 22;  Length 51;
  Best Local Similarity   88.2%;  Pred. No. 1.1e+03;
```

```
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 CATGTGATACCAGAAGG 17
          ||||||| | |||||||
Db     32 CATGTGACAGCAGAAGG 16


RESULT 6
AAQ12617
ID   AAQ12617 standard; DNA; 21 BP.
XX
AC   AAQ12617;
XX
DT   03-OCT-1991  (first entry)
XX
DE   Sense primer DRS18-1 for beta locus of TCR.
XX
KW   Restriction fragment length polymorphism; RFLP; T cell receptor; ss.
XX
OS   Homo sapiens.
XX
PN   WO9109623-A.
XX
PD   11-JUL-1991.
XX
PF   31-DEC-1990;   90WO-US07699.
XX
PR   01-MAY-1990;   90US-0517380.
PR   29-DEC-1989;   89US-0459065.
XX
PA   (CALY ) CALIF INST OF TECHN.
XX
PI   Urban JL,  Zaller DM,  Hood LE,  Beall SS,  Concannon P;
XX
DR   WPI; 1991-222662/30.
XX
PT   Diagnosis of auto:immune disease and antibodies for disease
PT   treatment - by detecting RFLP encoding variable region of T-cell
PT   antigen receptor beta-chain.
XX
PS   Claim 87; Page 131; 181pp; English.
XX
CC   The sequence is from the sense strand of the beta locus of the
CC   T-cell receptor.  The probe was used with a second oligo, DRS18-3,
CC   (AAQ12618) to amplify (by PCR) a non-coding region in a cosmid
CC   initially isolated by screening a total human cosmid library with
CC   various V region gene segments.  Sequencing of the region revealed
CC   an "anonymous" polymorphism (see AAQ12616).
CC   Presence of the polymorphism may indicate a predisposition to
CC   disease.
CC   See also AAQ12608-Q12623.
XX
SQ   Sequence 21 BP; 5 A; 4 C; 6 G; 6 T; 0 other;

  Query Match            68.0%;  Score 13.6;  DB 12;  Length 21;
  Best Local Similarity  80.0%;  Pred. No. 1.3e+03;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 CATGTGATACCAGAAGGTCA 20
          | | |||||| || ||||||
Db      1 CTTCTGATACTAGGAGGTCA 20


RESULT 7
AAS19393
ID   AAS19393 standard; DNA; 41 BP.
XX
AC   AAS19393;
XX
DT   26-MAR-2002  (first entry)
XX
DE   Human gap joining protein 10, associated probe.
XX
KW   Gap joining protein 10; human; cancer; human immunodeficiency virus;
KW   HIV; haemopathy; probe; ss.
XX
OS   Homo sapiens.
XX
PN   CN1315409-A.
XX
PD   03-OCT-2001.
XX
PF   28-MAR-2000; 2000CN-0115213.
XX
PR   28-MAR-2000; 2000CN-0115213.
XX
PA   (BODE-) BODE GENE DEV CO LTD SHANGHAI.
XX
PI   Mao Y,  Xie Y;
XX
DR   WPI; 2002-056322/08.
XX
PT   New polypeptide-human gap joining protein 10 for treating cancer,
PT   haemopathy, and human immunodeficiency virus infection, -
XX
PS   Example 6; Page 19(Disclosure); 32pp; Chinese.
XX
CC   The invention describes a new polypeptide-human gap joining protein 10,
CC   and the polynucleotide encoding it. A method for preparing the
CC   polypeptide by DNA recombination is also detailed. The polypeptide is
CC   useful for treating several diseases including cancer, haemopathy and
CC   human immunodeficiency virus (HIV) infection. The antagonist against the
CC   polypeptide and its therapeutic action, and the application of the
CC   polynucleotide for coding the new human gap joining protein 10 are
CC   disclosed. This sequence is probe, used to detect human gap joining
CC   protein 10, in the method described in example 6 of the invention.
XX
SQ   Sequence 41 BP; 13 A; 3 C; 9 G; 16 T; 0 other;

  Query Match            68.0%;  Score 13.6;  DB 24;  Length 41;
  Best Local Similarity  80.0%;  Pred. No. 1.4e+03;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 CATGTGATACCAGAAGGTCA 20
          ||| ||||| ||||| || |
Db     16 CATCTGATATCAGAATGTTA 35


RESULT 8
AAV70296
ID   AAV70296 standard; DNA; 53 BP.
XX
AC   AAV70296;
XX
DT   11-FEB-1999  (first entry)
XX
DE   Murine osteoprotegerin binding protein PCR primer 1581-73.
XX
KW   Osteoprotegerin binding protein; OPG binding protein; arthritis;
KW   osteoporosis; osteoclast maturation; bone disease; metastasis; ODAR;
KW   hypercalcaemia; osteoclast differentiation and activation receptor;
KW   Paget's disease; PCR primer; ss.
XX
OS   Synthetic.
OS   Mus sp.
XX
PN   WO9846751-A1.
XX
PD   22-OCT-1998.
XX
PF   15-APR-1998;   98WO-US07584.
XX
PR   30-MAR-1998;   98US-0052521.
PR   16-APR-1997;   97US-0842842.
PR   23-JUN-1997;   97US-0880855.
XX
```

```
PA   (AMGE-) AMGEN INC.
XX
PI   Boyle WJ;
XX
DR   WPI; 1998-594578/50.
XX
PT   Nucleic acid encoding osteoprotegrin binding protein - useful for,
PT   e.g. treating bone diseases by modulating osteoclast differentiation
PT   and for diagnosis
XX
PS   Example 6; Page 44; 47pp; English.
XX
CC   The present sequence represents a PCR primer for murine osteoprotegerin
CC   (OPG) binding protein, which is used in an example of the invention.
CC   Host cells transfected with vectors containing nucleic acid molecules
CC   encoding OPG binding protein are used to produce recombinant OPG binding
CC   protein. OPG binding protein is used in binding assays to determine
CC   osteoprotegrin (OG) in biological samples; to screen for specific
CC   binding agents (particularly agonists and antagonists, including
CC   intracellular proteins); to raise Ab (useful in immunoassays for
CC   detection of OPG binding protein) and to identify compounds that
CC   modulate binding of OPG binding protein to osteoclast differentiation
CC   and activation receptor (ODAR). The nucleic acid molecule encoding OPG
CC   binding protein can be used to detect OPG binding protein-encoding
CC   sequences, e.g. screening for related sequences, also to produce
CC   transgenic animal models, while complementary sequences are used for
CC   antisense regulation of OPG binding protein expression. Modulators of
CC   OPG binding protein, particularly soluble forms of OPG binding protein
CC   or Ab, are used to treat or prevent bone diseases, e.g. osteoporosis,
CC   bone loss caused by arthritis or metastases, hypercalcaemia, Paget's
CC   disease, periodontal disease, osteoporosis, loosening of prostheses,
CC   optionally in combination with agents that promote bone growth.
XX
SQ   Sequence 53 BP; 16 A; 18 C; 5 G; 14 T; 0 other;

  Query Match             68.0%;  Score 13.6;  DB 19;  Length 53;
  Best Local Similarity   80.0%;  Pred. No. 1.5e+03;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 CATGTGATACCAGAAGGTCA 20
           ||| ||| ||| |||| |||
Db       9 CATATGAAACCTGAAGCTCA 28


RESULT 9
ABN35867/c
ID   ABN35867 standard; DNA; 60 BP.
XX
AC   ABN35867;
XX
DT   15-JUL-2002  (first entry)
XX
DE   Human spliced transcript detection oligonucleotide SEQ ID NO:8615.
XX
KW   Human; mouse; rat; splice transcript; detection; RNA transcript;
KW   splice variant; transcriptome; oligonucleotide library; ss.
XX
OS   Homo sapiens.
XX
PN   WO200210449-A2.
XX
PD   07-FEB-2002.
XX
PF   20-JUL-2001; 2001WO-IB01903.
XX
PR   28-JUL-2000; 2000US-221607P.
PR   02-MAY-2001; 2001US-287724P.
XX
PA   (COMP-) COMPUGEN INC.
XX
PI   Shoshan A,  Wasserman A,  Mintz E,  Mintz L,  Faigler S;
XX
DR   WPI; 2002-257383/30.
XX
PT   New oligonucleotide libraries comprising oligonucleotides which
PT   selectively hybridize to mRNAs transcribed from a transcription unit of
PT   a genome, useful for detecting tissue-, pathology-, and
PT   developmental-specific genes  -
XX
PS   Example 1; SEQ ID 8615; 47pp; English.
XX
CC   The present invention describes oligonucleotide libraries for detecting
CC   messenger RNAs that populate a (sub-)transcriptome, where the
CC   (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC   transcription units that populate a genome. The library comprises
CC   several oligonucleotides, each capable of hybridising selectively to a
CC   set of messenger RNAs transcribed from a given transcription unit of
CC   the genome, which encodes one or more messenger RNA splice variants.
CC   The oligonucleotide libraries are useful for detecting mRNAs from a
CC   biological sample, in expression profiling studies, in qualitatively or
CC   quantitatively characterising the corresponding transcriptome, and in
CC   detecting RNA transcripts and splice variants of human or animal
CC   transcriptomes. The libraries may also be used as specialised mini
CC   libraries to detect transcripts of a sub-transcriptome under a
CC   particular biological or pathological state, and so allowing the
CC   detection of tissue- and pathology-specific genes such as those genes
CC   only expressed in specific tissue under a specific pathological
CC   condition; to detect developmental specific genes; and to detect RNA
CC   transcripts and splice variants of a transcriptome of a patient suffering
CC   from a particular disorder. ABN27253 to ABN59589 represent
CC   oligonucleotide sequences from rats, humans and mice, which are used in
CC   the exemplification of the present invention.
CC   N.B. The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 60 BP; 12 A; 23 C; 7 G; 18 T; 0 other;

  Query Match             68.0%;  Score 13.6;  DB 24;  Length 60;
  Best Local Similarity   80.0%;  Pred. No. 1.5e+03;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 CATGTGATACCAGAAGGTCA 20
           || ||||| |||||| || |
Db      53 CAAGTGATGCCAGAACGTGA 34


RESULT 10
ABK50854/c
ID   ABK50854 standard; DNA; 23 BP.
XX
AC   ABK50854;
XX
DT   30-JUL-2002  (first entry)
XX
DE   Lentinula edodes protein with laccase activity, degenerate primer #1.
XX
KW   Laccase; enzyme; primer; ss.
XX
OS   Lentinula edodes.
XX
FH   Key             Location/Qualifiers
FT   modified_base   3
FT                   /*tag= a
FT                   /mod_base= i
FT   modified_base   6
FT                   /*tag= b
FT                   /mod_base= i
FT   modified_base   21
FT                   /*tag= c
FT                   /mod_base= i
XX
PN   JP2002065282-A.
XX
PD   05-MAR-2002.
XX
```

```
PF   04-SEP-2000; 2000JP-0267473.
XX
PR   04-SEP-2000; 2000JP-0267473.
XX
PA   (IWAT-) IWATE KEN.
XX
DR   WPI; 2002-377662/41.
XX
PT   New proteins derived from Lentinula edodes for use in providing laccase
PT   enzyme activity  -
XX
PS   Example 3; Page 10; 30pp; Japanese.
XX
CC   The invention describes novel proteins derived from the species
CC   Lentinula edodes and having laccase enzyme activity. This sequence
CC   represents a degenerate PCR primer used in the invention.
XX
SQ   Sequence 23 BP; 4 A; 3 C; 4 G; 4 T; 8 other;

  Query Match             67.0%;  Score 13.4;  DB 24;  Length 23;
  Best Local Similarity   68.8%;  Pred. No. 1.6e+03;
  Matches   11;  Conservative    4;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       3 TGTGATACCAGAAGGT 18
           ::||:|||||:|| ||
Db      19 WRTGRTACCARAANGT 4


RESULT 11
ABN32733/c
ID   ABN32733 standard; DNA; 60 BP.
XX
AC   ABN32733;
XX
DT   15-JUL-2002  (first entry)
XX
DE   Human spliced transcript detection oligonucleotide SEQ ID NO:5481.
XX
KW   Human; mouse; rat; splice transcript; detection; RNA transcript;
KW   splice variant; transcriptome; oligonucleotide library; ss.
XX
OS   Homo sapiens.
XX
PN   WO200210449-A2.
XX
PD   07-FEB-2002.
XX
PF   20-JUL-2001; 2001WO-IB01903.
XX
PR   28-JUL-2000; 2000US-221607P.
PR   02-MAY-2001; 2001US-287724P.
XX
PA   (COMP-) COMPUGEN INC.
XX
PI   Shoshan A,  Wasserman A,  Mintz E,  Mintz L,  Faigler S;
XX
DR   WPI; 2002-257383/30.
XX
PT   New oligonucleotide libraries comprising oligonucleotides which
PT   selectively hybridize to mRNAs transcribed from a transcription unit of
PT   a genome, useful for detecting tissue-, pathology-, and
PT   developmental-specific genes  -
XX
PS   Example 1; SEQ ID 5481; 47pp; English.
XX
CC   The present invention describes oligonucleotide libraries for detecting
CC   messenger RNAs that populate a (sub-)transcriptome, where the
CC   (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC   transcription units that populate a genome. The library comprises
CC   several oligonucleotides, each capable of hybridising selectively to a
CC   set of messenger RNAs transcribed from a given transcription unit of
CC   the genome, which encodes one or more messenger RNA splice variants.
CC   The oligonucleotide libraries are useful for detecting mRNAs from a
CC   biological sample, in expression profiling studies, in qualitatively or
CC   quantitatively characterising the corresponding transcriptome, and in
CC   detecting RNA transcripts and splice variants of human or animal
CC   transcriptomes. The libraries may also be used as specialised mini
CC   libraries to detect transcripts of a sub-transcriptome under a
CC   particular biological or pathological state, and so allowing the
CC   detection of tissue- and pathology-specific genes such as those genes
CC   only expressed in specific tissue under a specific pathological
CC   condition; to detect developmental specific genes; and to detect RNA
CC   transcripts and splice variants of a transcriptome of a patient suffering
CC   from a particular disorder. ABN27253 to ABN59589 represent
CC   oligonucleotide sequences from rats, humans and mice, which are used in
CC   the exemplification of the present invention.
CC   N.B. The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 60 BP; 11 A; 15 C; 15 G; 19 T; 0 other;

  Query Match             67.0%;  Score 13.4;  DB 24;  Length 60;
  Best Local Similarity   93.3%;  Pred. No. 1.9e+03;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       6 GATACCAGAAGGTCA 20
           ||| ||||||||||
Db      52 GATCCCAGAAGGTCA 38


RESULT 12
ABN40118
ID   ABN40118 standard; DNA; 60 BP.
XX
AC   ABN40118;
XX
DT   15-JUL-2002  (first entry)
XX
DE   Human spliced transcript detection oligonucleotide SEQ ID NO:12866.
XX
KW   Human; mouse; rat; splice transcript; detection; RNA transcript;
KW   splice variant; transcriptome; oligonucleotide library; ss.
XX
OS   Homo sapiens.
XX
PN   WO200210449-A2.
XX
PD   07-FEB-2002.
XX
PF   20-JUL-2001; 2001WO-IB01903.
XX
PR   28-JUL-2000; 2000US-221607P.
PR   02-MAY-2001; 2001US-287724P.
XX
PA   (COMP-) COMPUGEN INC.
XX
PI   Shoshan A,  Wasserman A,  Mintz E,  Mintz L,  Faigler S;
XX
DR   WPI; 2002-257383/30.
XX
PT   New oligonucleotide libraries comprising oligonucleotides which
PT   selectively hybridize to mRNAs transcribed from a transcription unit of
PT   a genome, useful for detecting tissue-, pathology-, and
PT   developmental-specific genes  -
XX
PS   Example 1; SEQ ID 12866; 47pp; English.
XX
CC   The present invention describes oligonucleotide libraries for detecting
CC   messenger RNAs that populate a (sub-)transcriptome, where the
CC   (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC   transcription units that populate a genome. The library comprises
CC   several oligonucleotides, each capable of hybridising selectively to a
CC   set of messenger RNAs transcribed from a given transcription unit of
CC   the genome, which encodes one or more messenger RNA splice variants.
CC   The oligonucleotide libraries are useful for detecting mRNAs from a
```

CC biological sample, in expression profiling studies, in qualitatively or
CC quantitatively characterising the corresponding transcriptome, and in
CC detecting RNA transcripts and splice variants of human or animal
CC transcriptomes. The libraries may also be used as specialised mini
CC libraries to detect transcripts of a sub-transcriptome under a
CC particular biological or pathological state, and so allowing the
CC detection of tissue- and pathology-specific genes such as those genes
CC only expressed in specific tissue under a specific pathological
CC condition; to detect developmental specific genes; and to detect RNA
CC transcripts and splice variants of a transcriptome of a patient suffering
CC from a particular disorder. ABN27253 to ABN59589 represent
CC oligonucleotide sequences from rats, humans and mice, which are used in
CC the exemplification of the present invention.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 60 BP; 14 A; 19 C; 12 G; 15 T; 0 other;

Query Match 67.0%; Score 13.4; DB 24; Length 60;
Best Local Similarity 93.3%; Pred. No. 1.9e+03;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3 TGTGATACCAGAAGG 17
| ||||||||||||||
Db 33 TTTGATACCAGAAGG 47

RESULT 13
AAX08460
ID AAX08460 standard; DNA; 54 BP.
XX
AC AAX08460;
XX
DT 28-JUN-1999 (first entry)
XX
DE Primer for amplifying humanised green fluorescent protein clone.
XX
KW Green fluorescent protein; gfp; jellyfish; Aequorea victoria;
KW humanisation; reporter gene; substrate; cofactor; beta galactosidase;
KW firefly luciferase; alkaline phosphatase; PCR primer;
KW chloramphenicol acetyltransferase; CAT; beta glucuronidase; GUS; ss.
XX
OS Synthetic.
XX
PN WO9903997-A1.
XX
PD 28-JAN-1999.
XX
PF 16-JUL-1998; 98WO-US14692.
XX
PR 16-JUL-1997; 97US-0893327.
XX
PA (UYFL) UNIV FLORIDA.
XX
PI Hauswirth W, Muzyczka N, Zolotukhin S;
XX
DR WPI; 1999-132241/11.
XX
PT Humanised green fluorescent protein - used to measure gene
PT expression and identify transformed cells
XX
PS Example 5; Page 73; 152pp; English.
XX
CC Humanised green fluorescent protein (gfp) genes can be used to
CC identify transformed cells, to measure gene expression in vitro and
CC in vivo, to label specific cells in multicellular organisms (e.g. to
CC study cell lineage's), to label and locate fusion proteins, and to
CC study intracellular trafficking. Commonly used reporter genes include
CC beta-galactosidase, firefly luciferase, alkaline phosphatase;
CC chloramphenicol acetyltransferase (CAT), and beta glucuronidase
CC (GUS). However, these have limitations in their use. Frequently,
CC these reporter genes require the addition of a substrate and the
CC size of certain proteins means that the expression of reporter
CC fusion proteins can be difficult. The light stimulated GFP
CC fluorescence is species independent and does not require any
CC cofactors substrates or additional gene products from Aequorea
CC victoria an as the GFP genes have been humanised, they are
CC expressed at sufficient levels to be detectable in human cells,
CC unlike previous GFP proteins. This primer was used in conjunction
CC with a second primer (AAX08458) to produce a mutant GFP with
CC different excitation and emission spectra. This means that this
CC mutant GFP can be detected independently.
XX
SQ Sequence 54 BP; 15 A; 12 C; 15 G; 12 T; 0 other;

Query Match 66.0%; Score 13.2; DB 20; Length 54;
Best Local Similarity 83.3%; Pred. No. 2.4e+03;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CATGTGATACCAGAAGGT 18
|| | ||||||||||||
Db 35 CACGCCATACCAGAAGGT 52

RESULT 14
ABN54618
ID ABN54618 standard; DNA; 65 BP.
XX
AC ABN54618;
XX
DT 15-JUL-2002 (first entry)
XX
DE Mouse spliced transcript detection oligonucleotide SEQ ID NO:27366.
XX
KW Human; mouse; rat; splice transcript; detection; RNA transcript;
KW splice variant; transcriptome; oligonucleotide library; ss.
XX
OS Mus musculus.
XX
PN WO200210449-A2.
XX
PD 07-FEB-2002.
XX
PF 20-JUL-2001; 2001WO-IB01903.
XX
PR 28-JUL-2000; 2000US-221607P.
PR 02-MAY-2001; 2001US-287724P.
XX
PA (COMP-) COMPUGEN INC.
XX
PI Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
XX
DR WPI; 2002-257383/30.
XX
PT New oligonucleotide libraries comprising oligonucleotides which
PT selectively hybridize to mRNAs transcribed from a transcription unit of
PT a genome, useful for detecting tissue-, pathology-, and
PT developmental-specific genes -
XX
PS Example 1; SEQ ID 27366; 47pp; English.
XX
CC The present invention describes oligonucleotide libraries for detecting
CC messenger RNAs that populate a (sub-)transcriptome, where the
CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC transcription units that populate a genome. The library comprises
CC several oligonucleotides, each capable of hybridising selectively to a
CC set of messenger RNAs transcribed from a given transcription unit of
CC the genome, which encodes one or more messenger RNA splice variants.
CC The oligonucleotide libraries are useful for detecting mRNAs from a
CC biological sample, in expression profiling studies, in qualitatively or
CC quantitatively characterising the corresponding transcriptome, and in
CC detecting RNA transcripts and splice variants of human or animal
CC transcriptomes. The libraries may also be used as specialised mini
CC libraries to detect transcripts of a sub-transcriptome under a
CC particular biological or pathological state, and so allowing the

CC detection of tissue- and pathology-specific genes such as those genes
CC only expressed in specific tissue under a specific pathological
CC condition; to detect developmental specific genes; and to detect RNA
CC transcripts and splice variants of a transcriptome of a patient suffering
CC from a particular disorder. ABN27253 to ABN59589 represent
CC oligonucleotide sequences from rats, humans and mice, which are used in
CC the exemplification of the present invention.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 65 BP; 19 A; 12 C; 11 G; 23 T; 0 other;

Query Match 66.0%; Score 13.2; DB 24; Length 65;
Best Local Similarity 83.3%; Pred. No. 2.5e+03;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CATGTGATACCAGAAGGT 18
|||||| || |||||| |
Db 23 CATGTGTTATCAGAAGCT 40

RESULT 15
ABN54889/c
ID ABN54889 standard; DNA; 65 BP.
XX
AC ABN54889;
XX
DT 15-JUL-2002 (first entry)
XX
DE Mouse spliced transcript detection oligonucleotide SEQ ID NO:27637.
XX
KW Human; mouse; rat; splice transcript; detection; RNA transcript;
KW splice variant; transcriptome; oligonucleotide library; ss.
XX
OS Mus musculus.
XX
PN WO200210449-A2.
XX
PD 07-FEB-2002.
XX
PF 20-JUL-2001; 2001WO-IB01903.
XX
PR 28-JUL-2000; 2000US-221607P.
PR 02-MAY-2001; 2001US-287724P.
XX
PA (COMP-) COMPUGEN INC.
XX
PI Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
XX
DR WPI; 2002-257383/30.
XX
PT New oligonucleotide libraries comprising oligonucleotides which
PT selectively hybridize to mRNAs transcribed from a transcription unit of
PT a genome, useful for detecting tissue-, pathology-, and
PT developmental-specific genes -
XX
PS Example 1; SEQ ID 27637; 47pp; English.
XX
CC The present invention describes oligonucleotide libraries for detecting
CC messenger RNAs that populate a (sub-)transcriptome, where the
CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC transcription units that populate a genome. The library comprises
CC several oligonucleotides, each capable of hybridising selectively to a
CC set of messenger RNAs transcribed from a given transcription unit of
CC the genome, which encodes one or more messenger RNA splice variants.
CC The oligonucleotide libraries are useful for detecting mRNAs from a
CC biological sample, in expression profiling studies, in qualitatively or
CC quantitatively characterising the corresponding transcriptome, and in
CC detecting RNA transcripts and splice variants of human or animal
CC transcriptomes. The libraries may also be used as specialised mini
CC libraries to detect transcripts of a sub-transcriptome under a
CC particular biological or pathological state, and so allowing the

CC detection of tissue- and pathology-specific genes such as those genes
CC only expressed in specific tissue under a specific pathological
CC condition; to detect developmental specific genes; and to detect RNA
CC transcripts and splice variants of a transcriptome of a patient suffering
CC from a particular disorder. ABN27253 to ABN59589 represent
CC oligonucleotide sequences from rats, humans and mice, which are used in
CC the exemplification of the present invention.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 65 BP; 18 A; 16 C; 13 G; 18 T; 0 other;

Query Match 66.0%; Score 13.2; DB 24; Length 65;
Best Local Similarity 83.3%; Pred. No. 2.5e+03;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3 TGTGATACCAGAAGGTCA 20
||||| |||||||| | |
Db 46 TGTGAGACCAGAAGCTGA 29

Search completed: November 23, 2002, 07:03:57
Job time : 97.1 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 06:42:25 ; Search time 16.8 Seconds
(without alignments)
450.869 Million cell updates/sec

Title: US-09-296-264-27
Perfect score: 20
Sequence: 1 catgtgataccagaaggtca 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 335578 seqs, 189365133 residues

Total number of hits satisfying chosen parameters: 195614

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	20	100.0	30	10	US-09-986-632-30	Sequence 30, Appl
c	2	12.8	64.0	29	9	US-10-084-700-14	Sequence 14, Appl
	3	12.6	63.0	53	10	US-09-875-945-1	Sequence 1, Appli
	4	12.6	63.0	63	9	US-10-132-561-6	Sequence 6, Appli
c	5	12.2	61.0	25	10	US-09-899-980A-29	Sequence 29, Appl
c	6	12.2	61.0	79	10	US-09-864-761-26360	Sequence 26360, A
c	7	12	60.0	31	10	US-09-801-274-267	Sequence 267, App
	8	12	60.0	93	10	US-09-864-761-21619	Sequence 21619, A
c	9	12	60.0	99	10	US-09-864-761-31132	Sequence 31132, A
c	10	11.8	59.0	23	10	US-09-152-059-34	Sequence 34, Appl
c	11	11.8	59.0	31	10	US-09-801-274-7	Sequence 7, Appli
c	12	11.8	59.0	32	9	US-09-976-673-22	Sequence 22, Appl
	13	11.8	59.0	50	10	US-09-504-231A-2987	Sequence 2987, Ap
	14	11.8	59.0	50	10	US-09-274-553D-2987	Sequence 2987, Ap
	15	11.8	59.0	52	9	US-10-121-032-33	Sequence 33, Appl
	16	11.6	58.0	20	10	US-09-920-804-7	Sequence 7, Appli
	17	11.6	58.0	30	10	US-09-733-731-1	Sequence 7, Appli
	18	11.6	58.0	32	10	US-09-733-731-3	Sequence 3, Appli
	19	11.6	58.0	32	12	US-10-067-291-7	Sequence 7, Appli
	20	11.6	58.0	34	9	US-09-826-025-22	Sequence 22, Appl
	21	11.6	58.0	40	10	US-09-780-929-52	Sequence 52, Appl
c	22	11.6	58.0	96	10	US-09-864-761-19587	Sequence 19587, A
c	23	11.4	57.0	19	9	US-09-978-295A-562	Sequence 562, App
c	24	11.4	57.0	19	9	US-09-978-697-562	Sequence 562, App
	25	11.4	57.0	50	10	US-09-504-231A-3096	Sequence 3096, Ap
	26	11.4	57.0	50	10	US-09-274-553D-3096	Sequence 3096, Ap
	27	11.4	57.0	61	9	US-09-946-807-1357	Sequence 1357, Ap
	28	11.4	57.0	61	10	US-09-795-668-1357	Sequence 1357, Ap
	29	11.4	57.0	61	10	US-09-795-686-1357	Sequence 1357, Ap
	30	11.4	57.0	68	10	US-09-294-093B-2911	Sequence 2911, Ap
	31	11.4	57.0	76	10	US-09-864-761-30405	Sequence 30405, A
c	32	11.4	57.0	100	10	US-09-864-761-33290	Sequence 33290, A
c	33	11.2	56.0	18	10	US-09-216-393-145	Sequence 145, App
	34	11.2	56.0	19	10	US-09-729-674-236	Sequence 236, App
c	35	11.2	56.0	20	10	US-09-752-983-171	Sequence 171, App
c	36	11.2	56.0	23	8	US-08-812-393A-24	Sequence 24, Appl
	37	11.2	56.0	24	9	US-10-006-611-19	Sequence 19, Appl
	38	11.2	56.0	35	9	US-09-963-875-7	Sequence 7, Appli
c	39	11.2	56.0	49	10	US-09-005-243-97	Sequence 97, Appl
c	40	11.2	56.0	49	10	US-09-224-683-97	Sequence 97, Appl
	41	11.2	56.0	55	9	US-09-840-277-111	Sequence 111, App
	42	11.2	56.0	55	10	US-09-005-243-96	Sequence 96, Appl
	43	11.2	56.0	55	10	US-09-224-683-96	Sequence 96, Appl
	44	11.2	56.0	60	9	US-09-779-050A-36	Sequence 36, Appl
	45	11.2	56.0	81	10	US-09-998-598-943	Sequence 943, App

ALIGNMENTS

RESULT 1
US-09-986-632-30
; Sequence 30, Application US/09986632
; Patent No. US20020119944A1
; GENERAL INFORMATION:
; APPLICANT: AGUERA, Michelle
; TITLE OF INVENTION: Modulation of Ulip/CRMP activity for the prevention or
; TITLE OF INVENTION: treatment of myelin disorders
; FILE REFERENCE: P06974US01/BAS
; CURRENT APPLICATION NUMBER: US/09/986,632
; CURRENT FILING DATE: 2001-11-09
; PRIOR APPLICATION NUMBER: US 60/246,751
; PRIOR FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 30
; LENGTH: 30
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: antisense
US-09-986-632-30

Query Match 100.0%; Score 20; DB 10; Length 30;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 CATGTGATACCAGAAGGTCA 20
          ||||||||||||||||||||
Db      5 CATGTGATACCAGAAGGTCA 24
```

RESULT 2
US-10-084-700-14/c
; Sequence 14, Application US/10084700
; Patent No. US20020160403A1
; GENERAL INFORMATION:
; APPLICANT: Seeley, Todd
; TITLE OF INVENTION: huBUB3 GENE INVOLVED IN HUMAN CANCERS
; FILE REFERENCE: PP-01406.004/200130.438D1
; CURRENT APPLICATION NUMBER: US/10/084,700
; CURRENT FILING DATE: 2002-02-27

```
;  NUMBER OF SEQ ID NOS: 32
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 14
;   LENGTH: 29
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: PCR primer
US-10-084-700-14

  Query Match             64.0%;  Score 12.8;  DB 9;  Length 29;
  Best Local Similarity   87.5%;  Pred. No. 6.5e+02;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       5 TGATACCAGAAGGTCA 20
            |||| |||||||| ||
Db      29 TGATCCCAGAAGGGCA 14


RESULT 3
US-09-875-945-1
; Sequence 1, Application US/09875945
; Patent No. US20020098169A1
; GENERAL INFORMATION:
;  APPLICANT: METCON MEDICIN AB
;  APPLICANT:  SMITH, Ulf
;  TITLE OF INVENTION: No. US20020098169A1el sequences and their use
;  FILE REFERENCE: 45513MH
;  CURRENT APPLICATION NUMBER: US/09/875,945
;  CURRENT FILING DATE:  2001-06-08
;  PRIOR APPLICATION NUMBER: SE 0002189-9
;  PRIOR FILING DATE: 2000-06-09
;  PRIOR APPLICATION NUMBER: US 60/210,207
;  PRIOR FILING DATE: 2000-06-08
;  NUMBER OF SEQ ID NOS: 14
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 53
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-875-945-1

  Query Match             63.0%;  Score 12.6;  DB 10;  Length 53;
  Best Local Similarity   78.9%;  Pred. No. 9.3e+02;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 ATGTGATACCAGAAGGTCA 20
            |||||| ||   |||||||
Db       8 ATGTGAGACTTAAAGGTCA 26


RESULT 4
US-10-132-561-6
; Sequence 6, Application US/10132561
; Patent No. US20020168676A1
; GENERAL INFORMATION:
;  APPLICANT: No. US20020168676A1omi, Tsugunori
;  APPLICANT:  Hase, Tetsu
;  TITLE OF INVENTION: METHOD OF SYNTHESIZING NUCLEIC ACID
;  FILE REFERENCE: 201487/1020
;  CURRENT APPLICATION NUMBER: US/10/132,561
;  CURRENT FILING DATE:  2002-04-24
;  PRIOR APPLICATION NUMBER: US/09/530,061
;  PRIOR FILING DATE: 2000-04-21
;  PRIOR APPLICATION NUMBER: PCT/JP99/06213
;  PRIOR FILING DATE: 1999-11-08
;  PRIOR APPLICATION NUMBER: JP-1998-317476
;  PRIOR FILING DATE: 1998-11-09
;  NUMBER OF SEQ ID NOS: 29
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
;   LENGTH: 63
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:  Artificially
;   OTHER INFORMATION: synthesized primer sequence
US-10-132-561-6

  Query Match             63.0%;  Score 12.6;  DB 9;  Length 63;
  Best Local Similarity   78.9%;  Pred. No. 9.6e+02;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 ATGTGATACCAGAAGGTCA 20
            |||||| ||||||  || |
Db      24 ATGTGAAACCAGATCGTAA 42


RESULT 5
US-09-899-980A-29/c
; Sequence 29, Application US/09899980A
; Patent No. US20020058800A1
; GENERAL INFORMATION:
;  APPLICANT: Kingsbury, G.
;  APPLICANT:  Leiby, K.
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
;  TITLE OF INVENTION:  TREATMENT OF IMMUNE DISORDERS
;  FILE REFERENCE: 7853-158
;  CURRENT APPLICATION NUMBER: US/09/899,980A
;  CURRENT FILING DATE:  2001-07-06
;  PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/560,639
;  PRIOR FILING DATE: EARLIER FILING DATE: 2000-04-28
;  NUMBER OF SEQ ID NOS: 33
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 29
;   LENGTH: 25
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: 3' oligonucleotide
US-09-899-980A-29

  Query Match             61.0%;  Score 12.2;  DB 10;  Length 25;
  Best Local Similarity   82.4%;  Pred. No. 1.3e+03;
  Matches   14;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 CATGTGATACCAGAAGG 17
            | |||| || |||||||
Db      17 CTTGTGTTATCAGAAGG 1


RESULT 6
US-09-864-761-26360/c
; Sequence 26360, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
```

```
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 26360
;   LENGTH: 79
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AL021392.5
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3.3
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 7.9
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL =  19
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL =  14
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL =  16
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3.5
;   OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL =  5
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL =  11
;   OTHER INFORMATION: EST_HUMAN HIT: BE965632.2, EVALUE 9.40e-01
;   OTHER INFORMATION: NT HIT: AJ001056.1, EVALUE 4.00e-02
US-09-864-761-26360

  Query Match             61.0%;  Score 12.2;  DB 10;  Length 79;
  Best Local Similarity   82.4%;  Pred. No. 1.6e+03;
  Matches   14;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 ATGTGATACCAGAAGGT 18
           ||||| | |||||| ||
Db      28 ATGTGTTTCCAGAAAGT 12


RESULT 7
US-09-801-274-267/c
; Sequence 267, Application US/09801274
; Patent No. US20020032319A1
; GENERAL INFORMATION:
;  APPLICANT: Cargill, Michele
;  APPLICANT:  Ireland, James S.
;  APPLICANT:  Lander, Eric S.
;  TITLE OF INVENTION: HUMAN SINGLE NUCLEOTIDE POLYMORPHISMS
;  FILE REFERENCE: 2825.2009-001
;  CURRENT APPLICATION NUMBER: US/09/801,274
;  CURRENT FILING DATE:  2001-03-07
;  PRIOR APPLICATION NUMBER: US 60/187,510
;  PRIOR FILING DATE: 2000-03-07
;  PRIOR APPLICATION NUMBER: US 60/206,129
;  PRIOR FILING DATE: 2000-05-22
;  NUMBER OF SEQ ID NOS: 1802
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 267
;   LENGTH: 31
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-801-274-267

  Query Match             60.0%;  Score 12;  DB 10;  Length 31;
  Best Local Similarity   85.7%;  Pred. No. 1.7e+03;
  Matches   12;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       7 ATACCAGAAGGTCA 20
           |||||||||:| ||
Db      25 ATACCAGAARGCCA 12


RESULT 8
US-09-864-761-21619
; Sequence 21619, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 21619
;   LENGTH: 93
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AL109956.13
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.1
```

```
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.6
;   OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.1
;   OTHER INFORMATION: NT HIT: gi11545796, EVALUE 1.00e-45
;   OTHER INFORMATION: EST_HUMAN HIT: AU123146.1, EVALUE 2.00e-45
;   OTHER INFORMATION: SWISSPROT HIT: P46466, EVALUE 9.20e+00
US-09-864-761-21619

  Query Match             60.0%;  Score 12;  DB 10;  Length 93;
  Best Local Similarity   75.0%;  Pred. No. 2.1e+03;
  Matches   15;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      1 CATGTGATACCAGAAGGTCA 20
          || |||||  ||||| | ||
Db     17 CACGTGATTGCAGAAAGCCA 36


RESULT 9
US-09-864-761-31132/c
; Sequence 31132, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 31132
;   LENGTH: 99
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AC000004.1
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.73
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL =  18
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.7
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.79
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.79
;   OTHER INFORMATION: NT HIT: AL163303.2, EVALUE 5.00e-27
;   OTHER INFORMATION: EST_HUMAN HIT: F06752.1, EVALUE 7.00e-24
;   OTHER INFORMATION: SWISSPROT HIT: O08429, EVALUE 9.50e+00
US-09-864-761-31132

  Query Match             60.0%;  Score 12;  DB 10;  Length 99;
  Best Local Similarity   75.0%;  Pred. No. 2.2e+03;
  Matches   15;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      1 CATGTGATACCAGAAGGTCA 20
          |||| |||  ||| | ||||
Db     91 CATGAGATTGCAGCAAGTCA 72


RESULT 10
US-09-152-059-34/c
; Sequence 34, Application US/09152059
; Patent No. US20020068708A1
; GENERAL INFORMATION:
;  APPLICANT: WENGEL, JESPER
;  APPLICANT:  NIELSEN, POUL
;  TITLE OF INVENTION: OLIGONUCLEOTIDE ANALOGUES
;  FILE REFERENCE: 49165 (71994)
;  CURRENT APPLICATION NUMBER: US/09/152,059
;  CURRENT FILING DATE:  1998-09-11
;  PRIOR APPLICATION NUMBER: 60/058,541
;  PRIOR FILING DATE: 1997-09-12
;  PRIOR APPLICATION NUMBER: 60/068,293
;  PRIOR FILING DATE: 1997-12-19
;  PRIOR APPLICATION NUMBER: 60/071,682
;  PRIOR FILING DATE: 1998-01-16
;  PRIOR APPLICATION NUMBER: 60/076,591
;  PRIOR FILING DATE: 1998-03-03
;  PRIOR APPLICATION NUMBER: 60/083,507
;  PRIOR FILING DATE: 1998-04-29
;  PRIOR APPLICATION NUMBER: 60/088,309
;  PRIOR FILING DATE: 1998-06-05
;  PRIOR APPLICATION NUMBER: 60/094,355
;  PRIOR FILING DATE: 1998-07-28
;  NUMBER OF SEQ ID NOS: 146
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 34
;   LENGTH: 23
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Primer
US-09-152-059-34

  Query Match             59.0%;  Score 11.8;  DB 10;  Length 23;
  Best Local Similarity   86.7%;  Pred. No. 2.1e+03;
  Matches   13;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      3 TGTGATACCAGAAGG 17
          ||| |||||||||  |
Db     21 TGTCATACCAGAAAG 7


RESULT 11
US-09-801-274-7/c
; Sequence 7, Application US/09801274
; Patent No. US20020032319A1
; GENERAL INFORMATION:
;  APPLICANT: Cargill, Michele
;  APPLICANT:  Ireland, James S.
;  APPLICANT:  Lander, Eric S.
;  TITLE OF INVENTION: HUMAN SINGLE NUCLEOTIDE POLYMORPHISMS
```

```
;  FILE REFERENCE: 2825.2009-001
;  CURRENT APPLICATION NUMBER: US/09/801,274
;  CURRENT FILING DATE:  2001-03-07
;  PRIOR APPLICATION NUMBER: US 60/187,510
;  PRIOR FILING DATE: 2000-03-07
;  PRIOR APPLICATION NUMBER: US 60/206,129
;  PRIOR FILING DATE: 2000-05-22
;  NUMBER OF SEQ ID NOS: 1802
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7
;   LENGTH: 31
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-801-274-7

  Query Match             59.0%;  Score 11.8;  DB 10;  Length 31;
  Best Local Similarity   86.7%;  Pred. No. 2.2e+03;
  Matches   13;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       4 GTGATACCAGAAGGT 18
           |||||  ||||||||
Db      15 GTGATCGCAGAAGGT 1


RESULT 12
US-09-976-673-22/c
; Sequence 22, Application US/09976673
; Patent No. US20020160473A1
; GENERAL INFORMATION:
;  APPLICANT: Lukyanov, Sergey
;  APPLICANT:  Fradkov, Arcady
;  APPLICANT:  Labas, Yulii
;  APPLICANT:  Matz, Mikhail
;  APPLICANT:  Lukyanov, Konstantin
;  APPLICANT:  Gurskaya, Nadezda
;  TITLE OF INVENTION: FAR RED SHIFTED FLUORESCENT PROTEINS
;  FILE REFERENCE: CLON-028WO
;  CURRENT APPLICATION NUMBER: US/09/976,673
;  CURRENT FILING DATE:  2001-10-12
;  PRIOR APPLICATION NUMBER: 60/240,018
;  PRIOR FILING DATE: 2000-10-12
;  PRIOR APPLICATION NUMBER: 60,306,131
;  PRIOR FILING DATE: 2001-07-16
;  NUMBER OF SEQ ID NOS: 26
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 22
;   LENGTH: 32
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: primer
US-09-976-673-22

  Query Match             59.0%;  Score 11.8;  DB 9;  Length 32;
  Best Local Similarity   86.7%;  Pred. No. 2.2e+03;
  Matches   13;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       4 GTGATACCAGAAGGT 18
           ||||  |||||||| |
Db      20 GTGACACCAGAAGCT 6


RESULT 13
US-09-504-231A-2987
; Sequence 2987, Application US/09504231A
; Patent No. US20020013458A1
; GENERAL INFORMATION:
;  APPLICANT: Blatt, Lawrence
;  APPLICANT:  McSwiggen, James
;  APPLICANT:  Roberts, Beth
;  APPLICANT:  Pavco, Pamela
;  APPLICANT:  Macejak, Dennis
;  TITLE OF INVENTION: ENZYMATIC NUCLEIC ACID TREATMENT OF DISEASES OR CONDITIONS RELATE
;  TITLE OF INVENTION:  HEPATITIS C VIRUS INFECTION
;  FILE REFERENCE: rpi 247/282
;  CURRENT APPLICATION NUMBER: US/09/504,231A
;  CURRENT FILING DATE:  2000-02-15
;  PRIOR APPLICATION NUMBER: 09/274,553
;  PRIOR FILING DATE: 1999-03-23
;  PRIOR APPLICATION NUMBER: 09/257,608
;  PRIOR FILING DATE: 1999-02-24
;  PRIOR APPLICATION NUMBER: 60/100,842
;  PRIOR FILING DATE: 1998-09-18
;  PRIOR APPLICATION NUMBER: 60/083,217
;  PRIOR FILING DATE: 1998-04-27
;  NUMBER OF SEQ ID NOS: 3242
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2987
;   LENGTH: 50
;   TYPE: RNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:  Enzymatic Nucleic Acid Mo
US-09-504-231A-2987

  Query Match             59.0%;  Score 11.8;  DB 10;  Length 50;
  Best Local Similarity   73.3%;  Pred. No. 2.4e+03;
  Matches   11;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       6 GATACCAGAAGGTCA 20
           | :| |||||||:||
Db       1 GCUAUCAGAAGGUCA 15


RESULT 14
US-09-274-553D-2987
; Sequence 2987, Application US/09274553D
; Patent No. US20020082225A1
; GENERAL INFORMATION:
;  APPLICANT: Blatt, Lawrence
;  APPLICANT:  McSwiggen, James
;  APPLICANT:  Roberts, Beth
;  APPLICANT:  Pavco, Pamela
;  APPLICANT:  Macejak, Dennis
;  TITLE OF INVENTION: ENZYMATIC NUCLEIC ACID TREATMENT OF DISEASES OR CONDITIONS REL
;  TITLE OF INVENTION:  HEPATITIS C VIRUS INFECTION
;  FILE REFERENCE: rpi 247/282
;  CURRENT APPLICATION NUMBER: US/09/274,553D
;  CURRENT FILING DATE:  1999-03-23
;  PRIOR APPLICATION NUMBER: 09/257,608
;  PRIOR FILING DATE: 1999-02-24
;  PRIOR APPLICATION NUMBER: 60/100,842
;  PRIOR FILING DATE: 1998-09-18
;  PRIOR APPLICATION NUMBER: 60/083,217
;  PRIOR FILING DATE: 1998-04-27
;  NUMBER OF SEQ ID NOS: 3148
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2987
;   LENGTH: 50
;   TYPE: RNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:  Enzymatic Nucleic Acid Mo
US-09-274-553D-2987

  Query Match             59.0%;  Score 11.8;  DB 10;  Length 50;
  Best Local Similarity   73.3%;  Pred. No. 2.4e+03;
  Matches   11;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       6 GATACCAGAAGGTCA 20
           | :| |||||||:||
Db       1 GCUAUCAGAAGGUCA 15


RESULT 15
US-10-121-032-33
```

```
; Sequence 33, Application US/10121032
; Patent No. US20020155550A1
;   GENERAL INFORMATION:
;        APPLICANT: Bylina, Edward J.
;        TITLE OF INVENTION: GLYCOSIDASE ENZYMES
;        NUMBER OF SEQUENCES: 72
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Gray Cary Ware & Freidenrich LLP
;             STREET: 4365 Executive Drive, Suite 1600
;             CITY: San Diego
;             STATE: CA
;             COUNTRY: USA
;             ZIP: 92121
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Diskette
;             COMPUTER: IBM Compatible
;             OPERATING SYSTEM: Windows95
;             SOFTWARE: FastSEQ for Windows Version 2.0
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/10/121,032
;             FILING DATE: 09-Apr-2002
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US/09/134,078
;             FILING DATE: 13-AUG-1998
;             APPLICATION NUMBER: 08/949,026
;             FILING DATE: 10-OCT-1997
;             APPLICATION NUMBER: 60/056,916
;             FILING DATE: 06-DEC-1996
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Haile, Lisa A.
;             REGISTRATION NUMBER: 38,347
;             REFERENCE/DOCKET NUMBER: 09010/024002
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 858/677-1456
;             TELEFAX: 858/677-1465
;   INFORMATION FOR SEQ ID NO: 33:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 52 base pairs
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: Genomic DNA
;        SEQUENCE DESCRIPTION: SEQ ID NO: 33:
US-10-121-032-33

  Query Match             59.0%;  Score 11.8;  DB 9;  Length 52;
  Best Local Similarity   86.7%;  Pred. No. 2.4e+03;
  Matches   13;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       3 TGTGATACCAGAAGG 17
           | || ||||||||||
Db      31 TATGCTACCAGAAGG 45


Search completed: November 23, 2002, 07:10:43
Job time : 17.8 secs
```

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 26, 2002, 04:16:10 ; Search time 798.5 Seconds
(without alignments)
405.648 Million cell updates/sec

Title: US-09-296-264-27
Perfect score: 20
Sequence: 1 catgtgataccagaaggtca 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 357874

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	15.2	76.0	81	17	AZ370389	AZ370389 1M0121G19
	2	15.2	76.0	88	9	AA104892	AA104892 mo56b07.r
	3	14.2	71.0	72	17	AZ871178	AZ871178 2M0183K21
c	4	13.6	68.0	90	13	BI493128	BI493128 df97h12.y
	5	13.6	68.0	91	17	AZ388225	AZ388225 1M0148D08
	6	13.4	67.0	43	9	AA103211	AA103211 mo22a07.r
	7	13.4	67.0	70	9	AI938827	AI938827 sc60g06.y
c	8	13.2	66.0	38	17	AZ339876	AZ339876 1M0071J03
c	9	13	65.0	95	17	AZ782614	AZ782614 2M0023F15
c	10	12.8	64.0	33	14	R85322	R85322 yq25b08.s1
	11	12.8	64.0	48	14	N68498	N68498 za30a10.s1
c	12	12.8	64.0	52	17	BH853515	BH853515 SALK_0770
	13	12.8	64.0	72	17	TA320H08P	AL494364 T. brucei
c	14	12.8	64.0	93	9	AA874632	AA874632 vx81e04.r
c	15	12.8	64.0	99	17	AZ991984	AZ991984 2M0276P20
c	16	12.6	63.0	67	9	AA226274	AA226274 nc19d04.r
c	17	12.6	63.0	72	14	Z20794	Z20794 HSAAACWOG X
	18	12.6	63.0	76	14	BQ818585	BQ818585 1030071H0
	19	12.6	63.0	79	9	AA010006	AA010006 zi07c09.r
	20	12.6	63.0	79	9	AA259437	AA259437 va54a03.r
	21	12.4	62.0	53	10	AV963378	AV963378 AV963378
	22	12.4	62.0	67	17	CNS01WQ8	AL170729 Tetraodon
	23	12.4	62.0	82	17	AL763268	AL763268 Arabidops
c	24	12.4	62.0	83	9	AA022323	AA022323 mh85f02.r
	25	12.4	62.0	92	17	AZ575872	AZ575872 AST-T26B0
	26	12.4	62.0	94	9	AA999753	AA999753 os54h02.s
c	27	12.4	62.0	97	17	AZ790124	AZ790124 2M0038A02
	28	12.4	62.0	100	17	AZ694027	AZ694027 AST-2HLB4
	29	12.2	61.0	24	17	BH849325	BH849325 SALK_0695
	30	12.2	61.0	34	17	BH813677	BH813677 SALK_0645
	31	12.2	61.0	51	13	BI789980	BI789980 ic42h03.x
c	32	12.2	61.0	53	13	BM052387	BM052387 ic98a02.y
c	33	12.2	61.0	57	9	AA501216	AA501216 vh62f08.r
	34	12.2	61.0	62	17	AZ430539	AZ430539 1M0215L03
c	35	12.2	61.0	64	9	AI047729	AI047729 uh82e09.r
c	36	12.2	61.0	65	17	AZ661558	AZ661558 1M0540O03
c	37	12.2	61.0	67	9	AA662950	AA662950 nu96c01.s
c	38	12.2	61.0	73	13	BM570611	BM570611 ii04c07.y
c	39	12.2	61.0	77	12	BG315126	BG315126 PO1.0.75
c	40	12.2	61.0	77	13	BM507558	BM507558 ih41b12.y
c	41	12.2	61.0	78	13	BM054127	BM054127 id50h06.y
	42	12.2	61.0	79	9	AI219229	AI219229 qg13c08.x
	43	12.2	61.0	94	13	BI439925	BI439925 ic74c09.x
	44	12.2	61.0	94	14	BQ286441	BQ286441 ik30b05.y
c	45	12.2	61.0	95	9	AA200501	AA200501 mu34d03.r

ALIGNMENTS

RESULT 1
AZ370389/c
LOCUS AZ370389 81 bp DNA linear GSS 02-OCT-2000
DEFINITION 1M0121G19F Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC1M0121G19 F, DNA sequence.
ACCESSION AZ370389
VERSION AZ370389.1 GI:10484089
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 81)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D.,Weiss,R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00

```
            Plate: 0121  row: G  column: 19
            Seq primer: CGTTGTAAAACGACGGCCAGT
            Class: plasmid ends
            High quality sequence stop: 81.
FEATURES              Location/Qualifiers
     source           1. .81
                      /organism="Mus musculus"
                      /strain="C57BL/6J"
                      /db_xref="taxon:10090"
                      /clone="UUGC1M0121G19"
                      /clone_lib="Mouse 10kb plasmid UUGC1M library"
                      /sex="Male"
                      /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                      /note="Vector: PWD42nv; Purified genomic DNA from M.
                      musculus C57BL/6J (male) was obtained  from the Jackson
                      Laboratory Mouse DNA Resource
                      (http://www.jax.org/resources/documents/dnares/). The DNA
                      was hydrodynamically sheared by repeated passage through a
                      0.005 inch orifice at constant velocity. The sheared DNA
                      was blunt end-repaired with T4 DNA polymerase and T4
                      polynucleotide kinase. Adaptor oligonucleotides were
                      ligated to the blunt ends in high  molar excess. The
                      adaptored DNA was purified and size-selected for a 9.5 to
                      10.5 kb range using preparative agarose gel
                      electrophoresis. Vector DNA was prepared from a derivative
                      of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                      inducible derivative of plasmid R1. The vector was ligated
                      with adaptors complementary to the insert adaptors and
                      purified. The sheared, adaptored mouse DNA was annealed to
                      adaptored vector DNA, and transformed into
                      chemically-competent E. coli XL10-Gold (Stratagene) cells
                      and selected for ampicillin resistance."
BASE COUNT       10 a     11 c     25 g     35 t
ORIGIN

  Query Match             76.0%;  Score 15.2;  DB 17;  Length 81;
  Best Local Similarity   85.0%;  Pred. No. 3.7e+03;
  Matches   17;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 CATGTGATACCAGAAGGTCA 20
           ||||||| | |||||||| |
Db      39 CATGTGACATCAGAAGGTTA 20


RESULT 2
AA104892
LOCUS       AA104892                 88 bp    mRNA    linear   EST 29-OCT-1996
DEFINITION  mo56b07.r1 Life Tech mouse embryo 8 5dpc 10664019 Mus musculus cDNA
            clone IMAGE:557557 5' similar to TR:G1245118 G1245118 G PROTEIN
            COUPLED RECEPTOR. ;, mRNA sequence.
ACCESSION   AA104892
VERSION     AA104892.1  GI:1651060
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 88)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:338349
            Trace considered overall poor quality
            Possible reversed clone: similarity on wrong strand
            Seq primer: -28M13 rev1 from Amersham
            High quality sequence stop: 1.
FEATURES              Location/Qualifiers
     source           1. .88
                      /organism="Mus musculus"
                      /strain="C57BL/6J"
                      /db_xref="taxon:10090"
                      /clone="IMAGE:557557"
                      /clone_lib="Life Tech mouse embryo 8 5dpc 10664019"
                      /tissue_type="embryo"
                      /dev_stage="8.5dpc embryos"
                      /lab_host="DH10B"
                      /note="Organ: whole embryo; Vector: pCMV-SPORT2; Site_1:
                      SalI; Site_2: NotI; Cloned unidirectionally.  Primer:
                      Oligo dT.  8.5dpc embryos. pCMV-SPORT2 vector."
BASE COUNT       26 a     19 c     23 g     20 t
ORIGIN

  Query Match             76.0%;  Score 15.2;  DB 9;  Length 88;
  Best Local Similarity   85.0%;  Pred. No. 3.8e+03;
  Matches   17;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 CATGTGATACCAGAAGGTCA 20
           | |||  ||||||||||||
Db       7 CCTGTCTTACCAGAAGGTCA 26


RESULT 3
AZ871178
LOCUS       AZ871178                 72 bp    DNA     linear   GSS 21-FEB-2001
DEFINITION  2M0183K21R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC2M0183K21 R, DNA sequence.
ACCESSION   AZ871178
VERSION     AZ871178.1  GI:13077169
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 72)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0183  row: K  column: 21
            Seq primer: CACACAGGAAACAGCTATGACC
            Class: plasmid ends
            High quality sequence stop: 72.
FEATURES              Location/Qualifiers
     source           1. .72
                      /organism="Mus musculus"
                      /strain="C57BL/6J"
                      /db_xref="taxon:10090"
                      /clone="UUGC2M0183K21"
                      /clone_lib="Mouse 10kb plasmid UUGC1M library"
                      /sex="Male"
                      /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
```

/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptored DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptored mouse DNA was annealed to adaptored vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 24 a 19 c 13 g 16 t
ORIGIN

Query Match 71.0%; Score 14.2; DB 17; Length 72;
Best Local Similarity 84.2%; Pred. No. 1.1e+04;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy       2 ATGTGATACCAGAAGGTCA 20
           |||||||||| || ||| |
Db       4 ATGTGATACCCGACGGTAA 22
```

RESULT 4
BI493128/c
LOCUS BI493128 90 bp mRNA linear EST 28-AUG-2001
DEFINITION df97h12.y1 Morton Fetal Cochlea Homo sapiens cDNA clone IMAGE:2540782 5', mRNA sequence.
ACCESSION BI493128
VERSION BI493128.1 GI:15332472
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 90)
AUTHORS Robertson,N.G., Khetarpal,U., Gutierrez-Espeleta,G.A., Bieber,F.R. and Morton,C.C.
TITLE Isolation of novel and known genes from a human fetal cochlear cDNA library using subtractive hybridization and differential screening
JOURNAL Genomics 23, 42-50 (1994)
MEDLINE 95130111
COMMENT Contact: Morton, C. C.
Departments of Pathology and Obstetrics, Gynecology and Reproductive Biology
Brigham and Women's Hospital
75 Francis Street, Harvard Medical School, Boston, MA 02115, USA
Tel: 617 732 7980
Fax: 617 738 6996
Email: ccmorton@bics.bwh.harvard.edu
DNA sequencing and analyses were performed by National Institutes of Health Intramural Sequencing Center (NISC; see http://www.nisc.nih.gov).
This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Plate: LLAM6328 row: O column: 23
Seq primer: M13RP1 reverse primer (ABI).
FEATURES Location/Qualifiers
source 1. .90
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2540782"
/clone_lib="Morton Fetal Cochlea"
/tissue_type="cochlea"
/dev_stage="16-22 week fetus"
/lab_host="SOLR cells (kanamycin resistant)"
/note="Organ: ear; Vector: pBluescript SK-; Site_1: EcoRI; Site_2: XhoI; Reference: Genomics 23, 42-50 (1994) Cloned unidirectionally. Primer: Oligo dT. Fetal cochlea, normal. 37% of inserts <0.5 kb, 56% 0.5-1.0 kb, 7% >1 kb. Uni-ZAP XR Vector. Library constructed by N. Robertson, C. Morton. -5' adaptor sequence: 5' GAATTCGGCACGAG 3' -3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3'"
BASE COUNT 12 a 29 c 17 g 32 t
ORIGIN

Query Match 68.0%; Score 13.6; DB 13; Length 90;
Best Local Similarity 80.0%; Pred. No. 2.2e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
Qy       1 CATGTGATACCAGAAGGTCA 20
           |||  ||| |||| ||||||
Db      51 CATTGGATTCCAGCAGGTCA 32
```

RESULT 5
AZ388225
LOCUS AZ388225 91 bp DNA linear GSS 02-OCT-2000
DEFINITION 1M0148D08F Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC1M0148D08 F, DNA sequence.
ACCESSION AZ388225
VERSION AZ388225.1 GI:10501933
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 91)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D.,Weiss,R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0148 row: D column: 08
Seq primer: CGTTGTAAAACGACGGCCAGT
Class: plasmid ends
High quality sequence stop: 91.
FEATURES Location/Qualifiers
source 1. .91
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0148D08"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptored DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel

electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptored mouse DNA was annealed to adaptored vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."
BASE COUNT 20 a 15 c 29 g 27 t
ORIGIN

Query Match 68.0%; Score 13.6; DB 17; Length 91;
Best Local Similarity 80.0%; Pred. No. 2.2e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
Qy       1 CATGTGATACCAGAAGGTCA 20
           ||| ||| | | ||||||||
Db      45 CATATGAGAGCTGAAGGTCA 64
```

RESULT 6
AA103211
LOCUS AA103211 43 bp mRNA linear EST 29-OCT-1996
DEFINITION mo22a07.r1 Life Tech mouse embryo 13 5dpc 10666014 Mus musculus cDNA clone IMAGE:554292 5' similar to SW:HMG1_CHICK P36194 HIGH MOBILITY GROUP PROTEIN HMG1 ;, mRNA sequence.
ACCESSION AA103211
VERSION AA103211.1 GI:1649372
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 43)
AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.
TITLE The WashU-HHMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Marra M/Mouse EST Project
WashU-HHMI Mouse EST Project
Washington University School of MedicineP
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:335084
Trace considered overall poor quality
Possible reversed clone: similarity on wrong strand
Seq primer: -28M13 rev1 from Amersham
High quality sequence stop: 1.
FEATURES Location/Qualifiers
source 1. .43
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:554292"
/clone_lib="Life Tech mouse embryo 13 5dpc 10666014"
/tissue_type="embryo"
/dev_stage="13.5dpc embryos"
/lab_host="DH10B"
/note="Organ: whole embryo; Vector: pCMV-SPORT2; Site_1: SalI; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT. 13.5dpc embryos. pCMV-SPORT2 vector."
BASE COUNT 15 a 11 c 9 g 8 t
ORIGIN

Query Match 67.0%; Score 13.4; DB 9; Length 43;
Best Local Similarity 93.3%; Pred. No. 2.2e+04;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy       6 GATACCAGAAGGTCA 20
             ||||||||||| ||
Db      16 GATACCAGAAGGCCA 30
```

RESULT 7
AI938827
LOCUS AI938827 70 bp mRNA linear EST 30-NOV-2001
DEFINITION sc60g06.y1 Gm-c1016 Glycine max cDNA clone GENOME SYSTEMS CLONE ID: Gm-c1016-851 5' similar to SW:PRS7_PRUPE O64982 26S PROTEASE REGULATORY SUBUNIT 7 ;, mRNA sequence.
ACCESSION AI938827
VERSION AI938827.1 GI:5677697
KEYWORDS EST.
SOURCE soybean.
ORGANISM Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
REFERENCE 1 (bases 1 to 70)
AUTHORS Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Coryell,V., Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.
TITLE Public Soybean EST Project
JOURNAL Unpublished (1999)
COMMENT Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Trace considered overall poor quality
Possible reversed clone: similarity on wrong strand This clone is available through: ResGen, Invitrogen Corp. 2130 South Memorial Parkway Huntsville, AL 35801 For further information call: (800)-533-4363 or contact via email: ccu@resgen.com
High quality sequence stop: 1.
FEATURES Location/Qualifiers
source 1. .70
/organism="Glycine max"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-c1016-851"
/clone_lib="Gm-c1016"
/tissue_type="immature flowers of field grown plants"
/lab_host="XL10-Gold"
/note="Vector: pBluescript II XR; Site_1: EcoRI; Site_2: XhoI; This cDNA library was constructed from mRNA isolated from immature flowers of field grown plants. The cDNA library was prepared using the Stratagene pBluescript II XR library construction kit. Complementary DNA was synthesized from mRNA using a primer consisting of a poly (dT) sequence with a XhoI restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into XL10-Gold host cells. This library was constructed by Dr. Randy Shoemaker and Dr. John Erpelding."
BASE COUNT 24 a 16 c 13 g 17 t
ORIGIN

Query Match 67.0%; Score 13.4; DB 9; Length 70;
Best Local Similarity 93.3%; Pred. No. 2.6e+04;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy      6 GATACCAGAAGGTCA 20
          ||||||||||| |||
Db     27 GATACCAGAAGCTCA 41


RESULT 8
AZ339876/c
LOCUS       AZ339876                38 bp    DNA     linear   GSS 29-SEP-2000
DEFINITION  1M0071J03R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0071J03 R, DNA sequence.
ACCESSION   AZ339876
VERSION     AZ339876.1  GI:10414580
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 38)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0071  row: J  column: 03
            Seq primer: CACACAGGAAACAGCTATGACC
            Class: plasmid ends
            High quality sequence stop: 38.
FEATURES             Location/Qualifiers
     source          1. .38
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC1M0071J03"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
BASE COUNT       13 a      9 c      5 g     11 t
ORIGIN

  Query Match           66.0%;  Score 13.2;  DB 17;  Length 38;
  Best Local Similarity  83.3%;  Pred. No. 2.7e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;
```

```
Qy      1 CATGTGATACCAGAAGGT 18
          | |||||||||  |||||
Db     21 CCTGTGATACCCAAAGGT 4


RESULT 9
AZ782614/c
LOCUS       AZ782614                95 bp    DNA     linear   GSS 16-FEB-2001
DEFINITION  2M0023F15R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC2M0023F15 R, DNA sequence.
ACCESSION   AZ782614
VERSION     AZ782614.1  GI:12916513
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 95)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0023  row: F  column: 15
            Seq primer: CACACAGGAAACAGCTATGACC
            Class: plasmid ends
            High quality sequence stop: 95.
FEATURES             Location/Qualifiers
     source          1. .95
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC2M0023F15"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
BASE COUNT       26 a     18 c     17 g     34 t
ORIGIN

  Query Match           65.0%;  Score 13;  DB 17;  Length 95;
  Best Local Similarity  100.0%;  Pred. No. 4.3e+04;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy      6 GATACCAGAAGGT 18
          |||||||||||||
Db     68 GATACCAGAAGGT 56


RESULT 10
R85322/c
LOCUS       R85322                    33 bp    mRNA    linear   EST 22-JAN-1999
DEFINITION  yq25b08.s1 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone
            IMAGE:274527 3' similar to gb|M73048|HUMU3AAAA Human U3 small
            nuclear RNA (rRNA);, mRNA sequence.
ACCESSION   R85322
VERSION     R85322.1  GI:943728
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 33)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman
            ,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J.,
            Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston
            ,R., Williamson,A., Wohldmann,P. and Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Insert Size: 927
            High quality sequence starts: 1
            High quality sequence stops: 1
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Trace considered overall poor quality
            Insert Length: 927   Std Error: 0.00
            Seq primer: Promega -21m13
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .33
                     /organism="Homo sapiens"
                     /db_xref="GDB:3798645"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:274527"
                     /clone_lib="Soares fetal liver spleen 1NFLS"
                     /sex="male"
                     /dev_stage="20 week-post conception fetus"
                     /lab_host="DH10B (ampicillin resistant)"
                     /note="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia)
                     with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
                     1st strand cDNA was primed with a Pac I - oligo(dT) primer
                     [5' AACTGGAAGAATTAATTAAAGATCTTTTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Pac I and cloned into the Pac I
                     and Eco RI sites of the modified pT7T3 vector.  Library
                     went through one round of normalization. Library
                     constructed by Bento Soares and M.Fatima Bonaldo."
BASE COUNT        6 a      7 c     10 g      8 t      2 others
ORIGIN

  Query Match             64.0%;  Score 12.8;  DB 14;  Length 33;
  Best Local Similarity   87.5%;  Pred. No. 4e+04;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      5 TGATACCAGAAGGTCA 20
          |||| ||||||| |||
Db     31 TGATCCCAGAAGTTCA 16


RESULT 11
```

```
N68498
LOCUS       N68498                    48 bp    mRNA    linear   EST 13-MAR-1996
DEFINITION  za30a10.s1 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone
            IMAGE:294042 3' similar to gb:M65290 INTERLEUKIN-12 BETA CHAIN
            PRECURSOR (HUMAN);, mRNA sequence.
ACCESSION   N68498
VERSION     N68498.1  GI:1224659
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 48)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman
            ,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J.,
            Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston
            ,R., Williamson,A., Wohldmann,P. and Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Trace considered overall poor quality
            Seq primer: m13 -40 forward
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .48
                     /organism="Homo sapiens"
                     /db_xref="GDB:1238963"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:294042"
                     /clone_lib="Soares fetal liver spleen 1NFLS"
                     /sex="male"
                     /dev_stage="20 week-post conception fetus"
                     /lab_host="DH10B (ampicillin resistant)"
                     /note="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia)
                     with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
                     1st strand cDNA was primed with a Pac I - oligo(dT) primer
                     [5' AACTGGAAGAATTAATTAAAGATCTTTTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Pac I and cloned into the Pac I
                     and Eco RI sites of the modified pT7T3 vector.  Library
                     went through one round of normalization. Library
                     constructed by Bento Soares and M.Fatima Bonaldo."
BASE COUNT       12 a     10 c     16 g     10 t
ORIGIN

  Query Match             64.0%;  Score 12.8;  DB 14;  Length 48;
  Best Local Similarity   87.5%;  Pred. No. 4.5e+04;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      5 TGATACCAGAAGGTCA 20
          ||| ||||||||||
Db      1 TGAGCCCAGAAGGTCA 16


RESULT 12
BH853515/c
LOCUS       BH853515                  52 bp    DNA     linear   GSS 13-JUN-2002
DEFINITION  SALK_077078.16.20.x Arabidopsis thaliana TDNA insertion lines
            Arabidopsis thaliana genomic clone SALK_077078.16.20.x, DNA
            sequence.
ACCESSION   BH853515
VERSION     BH853515.1  GI:21424386
KEYWORDS    GSS.
SOURCE      thale cress.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
```

```
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 52)
  AUTHORS   Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R., Gadrinab
            ,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L., Shinn,P.
            , Zimmerman,J. and Ecker,J.R.
  TITLE     A Sequence-Indexed Library of Insertion Mutations in the
            Arabidopsis Genome
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Joseph R. Ecker
            Salk Institute Genomic Analysis Laboratory (SIGnAL)
            The Salk Institute for Biological Studies
            10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
            Tel: 858 453 4100 xi752
            Fax: 858 558 6379
            Email: ecker@salk.edu
            This is single pass sequence recovered from the left border of
            TDNA.
            Class: TDNA tagged.
FEATURES             Location/Qualifiers
     source          1. .52
                     /organism="Arabidopsis thaliana"
                     /strain="Columbia 0"
                     /db_xref="taxon:3702"
                     /clone="SALK_077078.16.20.x"
                     /clone_lib="Arabidopsis thaliana TDNA insertion lines"
                     /note="PCR was performed on Arabidopsis thaliana lines
                     each of which contains  one or more TDNA insertion
                     elements. The resultant fragment for each  line was
                     directly sequenced to determine the genomic sequence at
                     the  site of insertion. Details of the protocols used can
                     be found at  http://signal.salk.edu/tdna_protocols.html"
BASE COUNT       19 a      9 c      7 g     15 t      2 others
ORIGIN

  Query Match             64.0%;  Score 12.8;  DB 17;  Length 52;
  Best Local Similarity   77.8%;  Pred. No. 4.6e+04;
  Matches   14;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       3 TGTGATACCAGAAGGTCA 20
           |||||||  |  ||||||
Db      24 TGTGATANGANGAGGTCA 7


RESULT 13
TA320H08P
LOCUS       TA320H08P                72 bp    DNA     linear   GSS 13-DEC-2000
DEFINITION  T. brucei sheared genomic DNA clone 320h08, forward sequence,
            genomic survey sequence.
ACCESSION   AL494364
VERSION     AL494364.1  GI:11867879
KEYWORDS    GSS.
SOURCE      Trypanosoma brucei.
  ORGANISM  Trypanosoma brucei
            Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
            Trypanosoma.
REFERENCE   1  (bases 1 to 72)
  AUTHORS   Hall,N., Bowman,S., Lennard,N.J., Doggett,J., Atkin,R.,
            Chillingworth,C., Ormond,D., Harris,B., El-Sayed,N., Hou,L.,
            Melville,S.E., Rajandream,M.A. and Barrell,B.G.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-DEC-2000) Trypanosoma brucei genome sequencing
            project, Sanger Centre, The Wellcome Trust Genome Campus, Hinxton,
            Cambridge CB10 1SA, E-mail: barrell@sanger.ac.uk and
            nh1@sanger.ac.uk
COMMENT     Constructed at the Institute for Genomic Research (TIGR),
            Rockville, MD. Genomic DNA isolated from a cloned population of
            Trypanosoma brucei (TREU927/4 GUTat 10.1) was mechanically sheared
            to give a tight size distribution (
            4 kb).  The v + i method used for the library construction is
            described in detail in Smith, H. and Venter, J.C. (Making small
            insert libraries for whole genome shotgun sequencing projects.  In
            Genome Sequencing: A Practical Approach, eds. M. Vaudin and B.
            Barrell, Oxford University Press, 1999).
            Email: nelsayed@tigr.org
            Details of T. brucei sequencing at the Sanger Centre are available
            at http://www.sanger.ac.uk/Projects/T_brucei/.
FEATURES             Location/Qualifiers
     source          1. .72
                     /organism="Trypanosoma brucei"
                     /strain="TREU927"
                     /db_xref="taxon:5691"
                     /clone="320h08"
BASE COUNT       22 a     18 c     15 g     17 t
ORIGIN

  Query Match             64.0%;  Score 12.8;  DB 17;  Length 72;
  Best Local Similarity   87.5%;  Pred. No. 5e+04;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 CATGTGATACCAGAAG 16
           |||||||||  | ||||
Db      52 CATGTGATAGCGGAAG 67


RESULT 14
AA874632/c
LOCUS       AA874632                 93 bp    mRNA    linear   EST 19-MAR-1998
DEFINITION  vx81e04.r1 Soares_thymus_2NbMT Mus musculus cDNA clone
            IMAGE:1281630 5', mRNA sequence.
ACCESSION   AA874632
VERSION     AA874632.1  GI:2979321
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 93)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:673430
            Seq primer: -28m13 rev2 ET from Amersham
            High quality sequence stop: 51.
FEATURES             Location/Qualifiers
     source          1. .93
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:1281630"
                     /clone_lib="Soares_thymus_2NbMT"
                     /sex="male"
                     /tissue_type="Thymus"
                     /dev_stage="4 weeks"
                     /lab_host="DH10B"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                     was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCGTTTTTTTTTTTTTTTTTTTTTTTT
                     3']; double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. RNA
                     provided by Dr. Bertrand Jordan. Library went through two
```

rounds of normalization, and was constructed by Bento Soares and M.Fatima Bonaldo."
BASE COUNT 33 a 19 c 17 g 24 t
ORIGIN

Query Match 64.0%; Score 12.8; DB 9; Length 93;
Best Local Similarity 87.5%; Pred. No. 5.4e+04;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CATGTGATACCAGAAG 16
||| |||| |||||||
Db 30 CATCTGATGCCAGAAG 15

RESULT 15
AZ991984/c
LOCUS AZ991984 99 bp DNA linear GSS 27-APR-2001
DEFINITION 2M0276P20F Mouse 10kb plasmid UUGC2M library Mus musculus genomic clone UUGC2M0276P20 F, DNA sequence.
ACCESSION AZ991984
VERSION AZ991984.1 GI:13863211
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 99)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D.,Weiss,R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0276 row: P column: 20
Seq primer: CGTTGTAAAACGACGGCCAGT
Class: plasmid ends
High quality sequence stop: 99.
FEATURES Location/Qualifiers
source 1. .99
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC2M0276P20"
/clone_lib="Mouse 10kb plasmid UUGC2M library"
/sex="Female"
/lab_host="E. coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (female) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptored DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptored mouse DNA was annealed to adaptored vector. DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."
BASE COUNT 28 a 21 c 17 g 33 t
ORIGIN

Query Match 64.0%; Score 12.8; DB 17; Length 99;
Best Local Similarity 87.5%; Pred. No. 5.5e+04;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 2 ATGTGATACCAGAAGG 17
||||||||||| || |
Db 96 ATGTGATACCATAAAG 81

Search completed: November 26, 2002, 17:58:11
Job time : 807.5 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 3, 2002, 18:14:22 ; Search time 351.3 Seconds
(without alignments)
1656.863 Million cell updates/sec

Title: US-09-296-264-28
Perfect score: 20
Sequence: 1 ccaacaggcacagtacagca 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 995600

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		15.8	79.0	72	12	AF167288	AF167288 Cloning v
2	c	14	70.0	21	6	AX139384	AX139384 Sequence
3	c	14	70.0	21	6	BD010257	BD010257 Nucleic a
4		13.8	69.0	50	6	I88606	I88606 Sequence 2
5	c	13.8	69.0	51	6	AX116557	AX116557 Sequence
6		13.4	67.0	20	6	AR210628	AR210628 Sequence
7		13.4	67.0	20	6	AX047666	AX047666 Sequence
8		13.4	67.0	20	6	AX391295	AX391295 Sequence
9	c	13.4	67.0	54	10	S90689	S90689 V gamma 3,
10	c	13.4	67.0	63	9	HSU91086	U91086 Homo sapien
11	c	13.2	66.0	24	6	AX447198	AX447198 Sequence
12	c	13.2	66.0	35	6	AX179421	AX179421 Sequence
13	c	13.2	66.0	38	6	AX224713	AX224713 Sequence
14		13.2	66.0	51	6	AX203917	AX203917 Sequence
15	c	13	65.0	64	9	AF040159	AF040159 Homo sapi
16		12.8	64.0	23	6	AX055759	AX055759 Sequence
17		12.8	64.0	23	6	AX195470	AX195470 Sequence
18		12.8	64.0	23	6	AX403434	AX403434 Sequence
19		12.8	64.0	32	6	AR024108	AR024108 Sequence
20		12.8	64.0	32	6	BD001212	BD001212 Method an
21		12.8	64.0	32	6	BD001641	BD001641 Method an
22	c	12.8	64.0	48	6	A20891	A20891 fragment of
23	c	12.8	64.0	48	6	AR085995	AR085995 Sequence
24		12.8	64.0	48	6	AX026573	AX026573 Sequence
25		12.8	64.0	51	6	AX161689	AX161689 Sequence
26		12.8	64.0	51	6	AX161690	AX161690 Sequence
27		12.8	64.0	68	10	AF357455	AF357455 Mus muscu
28		12.8	64.0	91	17	HSJ227825	AJ227825 Homo sapi
29		12.8	64.0	95	10	MMV8IN38	Z12550 M.musculus
30		12.6	63.0	23	6	E51257	E51257 Disease tol
31		12.6	63.0	28	6	AR105938	AR105938 Sequence
32		12.6	63.0	29	6	AR083333	AR083333 Sequence
33	c	12.6	63.0	30	6	AR120438	AR120438 Sequence
34		12.6	63.0	35	6	E15066	E15066 Primer. 7/1
35		12.6	63.0	35	6	E15079	E15079 Primer. 7/1
36		12.6	63.0	35	6	E15104	E15104 Primer. 7/1
37		12.6	63.0	35	6	E27869	E27869 Method for
38	c	12.6	63.0	40	6	AR053595	AR053595 Sequence
39		12.6	63.0	41	6	AX327072	AX327072 Sequence
40		12.6	63.0	45	6	AX027547	AX027547 Sequence
41	c	12.6	63.0	51	6	AX161023	AX161023 Sequence
42	c	12.6	63.0	60	6	AR068214	AR068214 Sequence
43	c	12.6	63.0	60	6	AR076966	AR076966 Sequence
44	c	12.6	63.0	60	6	AR078799	AR078799 Sequence
45	c	12.6	63.0	60	6	AX136804	AX136804 Sequence

ALIGNMENTS

RESULT 1
AF167288
LOCUS AF167288 72 bp DNA linear SYN 06-JAN-2000
DEFINITION Cloning vector pJKS12 ZnO binding sequence.
ACCESSION AF167288
VERSION AF167288.1 GI:5733695
KEYWORDS .
SOURCE Cloning vector pJKS12.
ORGANISM Cloning vector pJKS12
artificial sequences; vectors.
REFERENCE 1 (bases 1 to 72)
AUTHORS Kjaergaard,K., Sorensen,J.K., Schembri,M.A. and Klemm,P.
TITLE Sequestration of zinc oxide by fimbrial designer chelators
JOURNAL Appl. Environ. Microbiol. 66 (1), 10-14 (2000)
MEDLINE 20087510

```
   PUBMED   10618196
REFERENCE   2  (bases 1 to 72)
  AUTHORS   Kjaergaard,K., Soerensen,J.K., Schembri,M.A. and Klemm,P.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-JUL-1999) Dept. of Microbiology, Technical University
            of Denmark, Lyngby 2800, Denmark
FEATURES             Location/Qualifiers
     source          1. .72
                     /organism="Cloning vector pJKS12"
                     /db_xref="taxon:101366"
                     /lab_host="Escherichia coli K-12"
     misc_feature    1. .72
                     /note="ZnO binding sequence enriched from screening of
                     random library in fimH"
BASE COUNT       26 a     20 c     16 g     10 t
ORIGIN

  Query Match             79.0%;  Score 15.8;  DB 12;  Length 72;
  Best Local Similarity   89.5%;  Pred. No. 7e+02;
  Matches   17;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      2 CAACAGGCACAGTACAGCA 20
          ||||| ||||| |||||||
Db     12 CAACAAGCACACTACAGCA 30


RESULT 2
AX139384/c
LOCUS       AX139384                 21 bp    DNA     linear   PAT 30-MAY-2001
DEFINITION  Sequence 1 from Patent EP1063301.
ACCESSION   AX139384
VERSION     AX139384.1  GI:14275060
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 21)
  AUTHORS   Suzuki,O., Ichihara,T., Shiohata,N. and Matsumura,Y.
  TITLE     Substrate carrying immobilised nucleic acid molecules
  JOURNAL   Patent: EP 1063301-A 1 27-DEC-2000;
            NISSHINBO INDUSTRIES, INC. (JP)
FEATURES             Location/Qualifiers
     source          1. .21
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="primer for PCR"
BASE COUNT        5 a      3 c      6 g      7 t
ORIGIN

  Query Match             70.0%;  Score 14;  DB 6;  Length 21;
  Best Local Similarity   100.0%;  Pred. No. 7.6e+03;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      3 AACAGGCACAGTAC 16
          ||||||||||||||
Db     20 AACAGGCACAGTAC 7


RESULT 3
BD010257/c
LOCUS       BD010257                 21 bp    DNA     linear   PAT 31-JAN-2002
DEFINITION  Nucleic acid-immobilized substrate.
ACCESSION   BD010257
VERSION     BD010257.1  GI:18638630
KEYWORDS    JP 2001066304-A/1.
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 21)
  AUTHORS   Suzuki,O., Ichihara,T., Shiohata,N. and Matsumura,Y.
  TITLE     Nucleic acid-immobilized substrate
  JOURNAL   Patent: JP 2001066304-A 1 16-MAR-2001;
            NISSHINBO IND INC
COMMENT     OS   Artificial Sequence
            PN   JP 2001066304-A/1
            PD   16-MAR-2001
            PF   15-JUN-2000 JP 2000179715
            PR
            PI   OSAMU SUZUKI,TATSUO ICHIHARA,NAMIKO SHIOHATA, PI   YOSHIYUKI
            MATSUMURA
            PC   G01N33/53,C12M1/00,C12M1/40,C12N15/09,G01N31/22, PC
            G01N33/566//C12Q1/68,
            PC   C12N15/00
            CC
            FH   Key             Location/Qualifiers
            FT   source          1. .21
            FT   /organism='Artificial Sequence'.
FEATURES             Location/Qualifiers
     source          1. .21
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
BASE COUNT        5 a      3 c      6 g      7 t
ORIGIN

  Query Match             70.0%;  Score 14;  DB 6;  Length 21;
  Best Local Similarity   100.0%;  Pred. No. 7.6e+03;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      3 AACAGGCACAGTAC 16
          ||||||||||||||
Db     20 AACAGGCACAGTAC 7


RESULT 4
I88606
LOCUS       I88606                   50 bp    DNA     linear   PAT 10-AUG-1998
DEFINITION  Sequence 2 from patent US 5718915.
ACCESSION   I88606
VERSION     I88606.1  GI:3408546
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Virtanen,J. and Virtanen,S.
  TITLE     Antiviral liposome having coupled target-binding moiety and
            hydrolytic enzyme
  JOURNAL   Patent: US 5718915-A 2 17-FEB-1998;
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="unknown"
BASE COUNT       19 a     15 c      8 g      7 t      1 others
ORIGIN

  Query Match             69.0%;  Score 13.8;  DB 6;  Length 50;
  Best Local Similarity   78.9%;  Pred. No. 9.9e+03;
  Matches   15;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 CAACAGGCACAGTACAGCA 20
          | |:|| ||||| ||||||
Db     18 CCAYAGTCACAGCACAGCA 36


RESULT 5
AX116557/c
LOCUS       AX116557                 51 bp    DNA     linear   PAT 11-MAY-2001
DEFINITION  Sequence 1680 from Patent WO0129262.
ACCESSION   AX116557
VERSION     AX116557.1  GI:14033499
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 51)
  AUTHORS   Picoult-Newburg,L. and Pohl,M.
```

```
  TITLE     Genotyping reagents, kits and methods of use thereof
  JOURNAL   Patent: WO 0129262-A 1680 26-APR-2001;
            Orchid BioSciences, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .51
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT       15 a      5 c     18 g     13 t
ORIGIN

  Query Match             69.0%;  Score 13.8;  DB 6;  Length 51;
  Best Local Similarity   88.2%;  Pred. No. 9.9e+03;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 CCAACAGGCACAGTACA 17
          ||||||||||||  | ||
Db     36 CCAACAGGCACATTCCA 20


RESULT 6
AR210628
LOCUS       AR210628                  20 bp    DNA     linear   PAT 20-JUN-2002
DEFINITION  Sequence 6 from patent US 6391311.
ACCESSION   AR210628
VERSION     AR210628.1  GI:21513406
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 20)
  AUTHORS   Ferrara,N. and Kuo,S.S.
  TITLE     Polypeptides having homology to vascular endothelial cell growth
            factor and bone morphogenetic protein 1
  JOURNAL   Patent: US 6391311-A 6 21-MAY-2002;
FEATURES             Location/Qualifiers
     source          1. .20
                     /organism="unknown"
BASE COUNT        7 a      8 c      3 g      2 t
ORIGIN

  Query Match             67.0%;  Score 13.4;  DB 6;  Length 20;
  Best Local Similarity   93.3%;  Pred. No. 1.7e+04;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      2 CAACAGGCACAGTAC 16
          |||||||||||| |
Db      2 CAACAGGCACAGTTC 16


RESULT 7
AX047666
LOCUS       AX047666                  20 bp    DNA     linear   PAT 15-DEC-2000
DEFINITION  Sequence 25 from Patent WO0070050.
ACCESSION   AX047666
VERSION     AX047666.1  GI:11876708
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 20)
  AUTHORS   Baker,K.P., Chen,J., Ferrara,N., Fong,S., Goddard,A., Gurney,A.L.,
            Hillan,K.J., Kuo,S.S., Tumas,D. and Wood,W.I.
  TITLE     Compositions and methods for the treatment of immune related
            diseases
  JOURNAL   Patent: WO 0070050-A 25 23-NOV-2000;
            Genentech, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .20
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="Cloning Oligo"
BASE COUNT        7 a      8 c      3 g      2 t
ORIGIN

  Query Match             67.0%;  Score 13.4;  DB 6;  Length 20;
  Best Local Similarity   93.3%;  Pred. No. 1.7e+04;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      2 CAACAGGCACAGTAC 16
          |||||||||||| |
Db      2 CAACAGGCACAGTTC 16


RESULT 8
AX391295
LOCUS       AX391295                  20 bp    DNA     linear   PAT 23-MAR-2002
DEFINITION  Sequence 38 from Patent WO0073445.
ACCESSION   AX391295
VERSION     AX391295.1  GI:19699952
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Ferrara,N., Williams,P.M., Baker,K.P., Ashkenazi,A.J., Goddard,A.,
            Godowski,P.J., Gurney,A.L., Kuo,S.S., Mark,M.R., Marsters,S.A.,
            Pitti,R.M., Wood,W.I., Gerber,H., Gerritsen,M.E., Paoni,N.F. and
            Watanabe,C.K.
  TITLE     Promotion or inhibition of angiogenesis and cardiovascularization
  JOURNAL   Patent: WO 0073445-A 38 07-DEC-2000;
            Genentech Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .20
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="Synthetic oligonucleotide probe"
BASE COUNT        7 a      8 c      3 g      2 t
ORIGIN

  Query Match             67.0%;  Score 13.4;  DB 6;  Length 20;
  Best Local Similarity   93.3%;  Pred. No. 1.7e+04;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      2 CAACAGGCACAGTAC 16
          |||||||||||| |
Db      2 CAACAGGCACAGTTC 16


RESULT 9
S90689/c
LOCUS       S90689                    54 bp    DNA     linear   ROD 07-MAY-1993
DEFINITION  V gamma 3, TcR-TcR C gamma 1 chain V-J region [mice, FF3 cells,
            Genomic, 54 nt].
ACCESSION   S90689
VERSION     S90689.1  GI:246310
KEYWORDS    .
SOURCE      Mus sp. FF3 cells.
  ORGANISM  Mus sp.
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 54)
  AUTHORS   Ezquerra,A., Wilde,D.B., McConnell,T.J., Sturmhofel,K., Valas,R.B.,
            Shevach,E.M. and Coligan,J.E.
  TITLE     Mouse autoreactive gamma/delta T cells. II. Molecular
            characterization of the T cell receptor
  JOURNAL   Eur. J. Immunol. 22 (2), 491-498 (1992)
  MEDLINE   92164730
   PUBMED   1311262
  REMARK    GenBank staff at the National Library of Medicine created this
            entry [NCBI gibbsq 90689] from the original journal article.
            This sequence comes from Fig. 5.
FEATURES             Location/Qualifiers
     source          1. .54
                     /organism="Mus sp."
                     /db_xref="taxon:10095"
     gene            1. .54
```

```
                     /partial
                     /gene="V<gamma>3, TcR"
                     /note="TcR C gamma 1 chain V-J region"
     CDS             1. .54
                     /partial
                     /gene="V<gamma>3, TcR"
                     /note="hybridoma FF3; This sequence comes from Fig. 5"
                     /codon_start=1
                     /product="TcR C gamma 1 chain V-J region"
                     /protein_id="AAB21558.1"
                     /db_xref="GI:246311"
                     /translation="ATYYCACWDSSGFHKVFA"
BASE COUNT       11 a     13 c     14 g     16 t
ORIGIN

  Query Match             67.0%;  Score 13.4;  DB 10;  Length 54;
  Best Local Similarity   93.3%;  Pred. No. 1.7e+04;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 CCAACAGGCACAGTA 15
          ||| ||||||||||||
Db     24 CCAGCAGGCACAGTA 10


RESULT 10
HSU91086/c
LOCUS       HSU91086                  63 bp    mRNA    linear   PRI 08-JUL-1997
DEFINITION  Homo sapiens T-cell receptor delta chain (TCRDV2J2) mRNA, partial
            cds.
ACCESSION   U91086
VERSION     U91086.1  GI:2239579
KEYWORDS    .
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 63)
  AUTHORS   Holtmeier,W., Witthoft,T., Hennemann,A., Winter,H.S. and
            Kagnoff,M.F.
  TITLE     The TCR-delta repertoire in human intestine undergoes
            characteristic changes during fetal to adult development
  JOURNAL   J. Immunol. 158 (12), 5632-5641 (1997)
  MEDLINE   97334214
   PUBMED   9190911
REFERENCE   2  (bases 1 to 63)
  AUTHORS   Holtmeier,W., Witthoeft,T., Hennemann,A., Harland,S.W. and
            Kagnoff,M.F.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-FEB-1997) Department of Medicine, University of
            Frankfurt, Theodor-Stern Kai #7, Frankfurt 60590, Germany
FEATURES             Location/Qualifiers
     source          1. .63
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="14"
                     /clone="FE1528"
                     /tissue_type="colon"
                     /dev_stage="fetus; 20 week gestation"
     gene            <1. .>63
                     /gene="TCRDV2J2"
     CDS             <1. .>63
                     /gene="TCRDV2J2"
                     /note="rearranged; contains CDR3 domain, length 09 amino
                     acids (calculation according to Rock,E.
                     J.Exp.Med.179:323-328,1994)"
                     /codon_start=1
                     /product="T-cell receptor delta chain"
                     /protein_id="AAB62446.1"
                     /db_xref="GI:2239580"
                     /translation="ERDEGSYYCACGGYALTAQLF"
     V_region        <1. .>63
                     /gene="TCRDV2J2"
BASE COUNT       15 a     13 c     19 g     16 t
ORIGIN

  Query Match             67.0%;  Score 13.4;  DB 9;  Length 63;
  Best Local Similarity   93.3%;  Pred. No. 1.7e+04;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 CCAACAGGCACAGTA 15
          || ||||||||||||
Db     36 CCCACAGGCACAGTA 22


RESULT 11
AX447198/c
LOCUS       AX447198                  24 bp    DNA     linear   PAT 05-JUL-2002
DEFINITION  Sequence 3653 from Patent WO0216649.
ACCESSION   AX447198
VERSION     AX447198.1  GI:21696097
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Gunderson,K.
  TITLE     Probes and decoder oligonucleotides
  JOURNAL   Patent: WO 0216649-A 3653 28-FEB-2002;
            Illumina, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .24
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="Computer Generated Probe Sequence."
BASE COUNT        2 a      7 c      7 g      8 t
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 6;  Length 24;
  Best Local Similarity   83.3%;  Pred. No. 2.2e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      3 AACAGGCACAGTACAGCA 20
          ||||| ||| |||| |||
Db     24 AACAGACACCGTACTGCA 7


RESULT 12
AX179421/c
LOCUS       AX179421                  35 bp    DNA     linear   PAT 04-JUL-2001
DEFINITION  Sequence 16 from Patent WO0131037.
ACCESSION   AX179421
VERSION     AX179421.1  GI:14599069
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 35)
  AUTHORS   Ben-Nun,A., kerlero de Rosbo,N. and Sappler,G.P.
  TITLE     Synthetic human genes and polypeptides and their use in the
            treatment of autoimmune diseases
  JOURNAL   Patent: WO 0131037-A 16 03-MAY-2001;
            YEDA RESEARCH AND DEVELOPMENT CO. LTD. (IL)
FEATURES             Location/Qualifiers
     source          1. .35
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
BASE COUNT        5 a      8 c     12 g     10 t
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 6;  Length 35;
  Best Local Similarity   83.3%;  Pred. No. 2.2e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 CAACAGGCACAGTACAGC 19
          | |||| ||||| |||||
Db     31 CCACAGCCACAGAACAGC 14
```

```
RESULT 13
AX224713/c
LOCUS       AX224713                38 bp    DNA     linear   PAT 10-SEP-2001
DEFINITION  Sequence 8 from Patent WO0160985.
ACCESSION   AX224713
VERSION     AX224713.1  GI:15554826
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 38)
  AUTHORS   Pizzorno,G., Cao,D. and Zhang,D.
  TITLE     Compositions, methods and kits relating to uridine phosphorylase
            gene mutations
  JOURNAL   Patent: WO 0160985-A 8 23-AUG-2001;
            YALE UNIVERSITY (US)
FEATURES             Location/Qualifiers
     source          1. .38
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="forward primer"
BASE COUNT        8 a      14 c      6 g      10 t
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 6;  Length 38;
  Best Local Similarity   83.3%;  Pred. No. 2.2e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 CAACAGGCACAGTACAGC 19
           || | |||| ||||||||
Db      18 CACCGGGCATAGTACAGC 1


RESULT 14
AX203917
LOCUS       AX203917                51 bp    DNA     linear   PAT 30-AUG-2001
DEFINITION  Sequence 23 from Patent WO0148245.
ACCESSION   AX203917
VERSION     AX203917.1  GI:15393372
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 51)
  AUTHORS   Shimkets,R.A. and Leach,M.
  TITLE     Nucleic acids containing single nucleotide polymorphisms and
            methods of use thereof
  JOURNAL   Patent: WO 0148245-A 23 05-JUL-2001;
            Curagen Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .51
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     variation       26
                     /note="single nucleotide polymorphism
                     Accession number cg1395871"
BASE COUNT       18 a      14 c      16 g       3 t
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 6;  Length 51;
  Best Local Similarity   83.3%;  Pred. No. 2.2e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 CCAACAGGCACAGTACAG 18
           ||| ||||||||| | ||
Db       9 CCAGCAGGCACAGGAAAG 26


RESULT 15
AF040159/c
LOCUS       AF040159                64 bp    mRNA    linear   PRI 11-FEB-1998
DEFINITION  Homo sapiens multiple sclerosis patient, T cell receptor delta
            chain, rearranged Vdelta2-Jdelta3 region, mRNA, partial cds.
ACCESSION   AF040159
VERSION     AF040159.1  GI:2853630
KEYWORDS    .
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 64)
  AUTHORS   Liedtke,W., Meyer,G., Faustmann,P.M., Warnatz,H. and Raine,C.S.
  TITLE     Clonal expansion and decreased occurrence of peripheral blood gamma
            delta T cells of the V delta 2J delta 3 lineage in multiple
            sclerosis patients
  JOURNAL   Int. Immunol. 9 (7), 1031-1041 (1997)
  MEDLINE   97380322
   PUBMED   9237112
REFERENCE   2  (bases 1 to 64)
  AUTHORS   Liedtke,W., Meyer,G., Faustmann,P.M., Warnatz,H. and Raine,C.S.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-DEC-1997) Molecular Genetics, The Rockefeller
            University, Howard Hughes Medical Institute, Box 305, 1230 York
            Avenue, New York, NY 10021, USA
FEATURES             Location/Qualifiers
     source          1. .64
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /cell_type="peripheral blood gamma/delta T cells"
                     /rearranged
                     /note="multiple sclerosis patient from a copolymer I
                     clinical trial"
     CDS             <1. .>64
                     /note="Vdelta2-Jdelta3 rearrangement"
                     /codon_start=1
                     /product="T cell receptor delta chain"
                     /protein_id="AAC02435.1"
                     /db_xref="GI:2853631"
                     /translation="ERDEGSYYCACFPLGDTHSWD"
BASE COUNT       13 a      15 c      20 g      16 t
ORIGIN

  Query Match             65.0%;  Score 13;  DB 9;  Length 64;
  Best Local Similarity   100.0%;  Pred. No. 2.9e+04;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3 AACAGGCACAGTA 15
           |||||||||||||
Db      34 AACAGGCACAGTA 22


Search completed: December  3, 2002, 22:23:45
Job time : 358.3 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 06:29:45 ; Search time 94.1 Seconds
(without alignments)
478.639 Million cell updates/sec

Title: US-09-296-264-28
Perfect score: 20
Sequence: 1 ccaacaggcacagtacagca 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 2390332

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		20	100.0	20	21	AAZ31458	Human neuropilin m
c 2		16.8	84.0	51	22	AAL33140	Human SNP oligonuc
c 3		15.8	79.0	51	22	AAL33141	Human SNP oligonuc
c 4		15.2	76.0	51	22	AAL33142	Human SNP oligonuc
c 5		14.8	74.0	24	24	ABN80383	Oligonucleotide pr
c 6		14.2	71.0	52	22	AAH42148	HRE element from t
c 7		14.2	71.0	65	24	ABN52993	Mouse spliced tran
8		14	70.0	20	24	ABL45353	Human chromosome 2
9		14	70.0	21	22	AAF74881	Chicken neuropilin
c 10		14	70.0	21	22	AAC91777	Escherichia coli O
c 11		13.8	69.0	22	22	AAF98589	Human kinase marke
c 12		13.8	69.0	29	21	AAA04344	Polymorphic fragme
c 13		13.8	69.0	51	22	AAH38884	Human SNP flanking
c 14		13.8	69.0	60	24	ABN34175	Human spliced tran
15		13.8	69.0	60	24	ABN41646	Human spliced tran
16		13.8	69.0	65	24	ABN50982	Mouse spliced tran
c 17		13.6	68.0	51	22	AAL32407	Human SNP oligonuc
18		13.6	68.0	60	16	AAQ82729	5' primer B1789CC
19		13.6	68.0	60	24	ABN58531	Human spliced tran
20		13.6	68.0	60	24	ABN58793	Human spliced tran
c 21		13.6	68.0	65	24	ABN28500	Rat spliced transc
22		13.6	68.0	65	24	ABN29623	Rat spliced transc
23		13.6	68.0	65	24	ABN53043	Mouse spliced tran
c 24		13.6	68.0	72	24	AAD30974	Arabidopsis thalia
25		13.6	68.0	81	16	AAQ82738	5' PCR primer C348
26		13.6	68.0	82	16	AAQ82734	5' PCR primer C335
27		13.6	68.0	82	16	AAQ82736	5' PCR primer C348
28		13.6	68.0	91	16	AAQ82732	5' PCR primer C335
29		13.4	67.0	20	20	AAZ23688	Human VEGF-E DNA p
30		13.4	67.0	20	20	AAZ34300	Human PRO200 hybri
31		13.4	67.0	20	21	AAC78884	Human PRO200 hybri
32		13.4	67.0	20	21	AAA88526	PRO200 (VEGF-E) DN
33		13.4	67.0	20	21	AAC58400	Human PRO200 (UNQ1
34		13.4	67.0	20	21	AAA77625	Human PRO713 oligo
35		13.4	67.0	20	22	AAC90583	Human PRO polynucl
36		13.4	67.0	20	22	AAC88977	Human PRO200 codin
37		13.4	67.0	20	22	AAC97408	Human PRO200 hybri
c 38		13.4	67.0	65	24	ABN53526	Mouse spliced tran
c 39		13.2	66.0	24	24	ABQ03646	Oligonucleotide ad
c 40		13.2	66.0	35	22	AAS03833	Synthetic gene shP
c 41		13.2	66.0	38	23	ABK51243	Human uridine phos
42		13.2	66.0	50	22	AAL31538	Human SNP oligonuc
43		13.2	66.0	51	22	AAH79408	Human DNA containi
c 44		13.2	66.0	60	24	ABN35448	Human spliced tran
c 45		13.2	66.0	60	24	ABN37997	Human spliced tran

ALIGNMENTS

RESULT 1
AAZ31458
ID AAZ31458 standard; DNA; 20 BP.
XX
AC AAZ31458;
XX
DT 07-FEB-2000 (first entry)
XX
DE Human neuropilin mRNA specific antisense oligo GTI3629.
XX
KW Neuropilin; human; growth; metastasis; tumor; neovascularisation;
KW cancer; papilloma; diabetic retinopathy; antisense; ss.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9955855-A2.
XX
PD 04-NOV-1999.
XX
PF 23-APR-1999; 99WO-CA00324.
XX
PR 23-APR-1998; 98US-0082791.
XX
PA (GENE-) GENESENSE TECHNOLOGIES INC.
XX
PI Wright JA, Young AH, Lee YS;
XX
DR WPI; 2000-023357/02.
XX
PT Antisense oligonucleotides that inhibit neuropilin expression, useful
PT for treating cancer -

XX
PS Claim 4; Page 17; 57pp; English.
XX
CC Sequences AAZ31431-460 represent antisense oligonucleotides which
CC inhibit human neuropilin expression. The antisense oligonucleotides can
CC be used to inhibit the growth or metastasis of a mammalian tumor and
CC inhibit neovascularisation. The oligonucleotides may be used to treat
CC various forms of cancers or tumors, such as sarcomas, melanomas,
CC adenomas, carcinomas of solid tissue, hypoxic tumors, squamous cell
CC carcinomas of the mouth, throat, larynx and lung, genitourinary cancers
CC such as cervical and bladder cancer, hematopoietic cancers, colon cancer,
CC breast cancer, pancreatic cancer, renal cancer, brain cancer, skin
CC cancer, liver cancer, head and neck cancers, and nervous system cancers,
CC as well as benign lesions such as papillomas. The methods may be used to
CC treat neovascularisation disorders such as diabetic retinopathy, and
CC retinopathy of prematurity and age related macular degeneration.
XX
SQ Sequence 20 BP; 8 A; 7 C; 4 G; 1 T; 0 other;

Query Match 100.0%; Score 20; DB 21; Length 20;
Best Local Similarity 100.0%; Pred. No. 3.3;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 CCAACAGGCACAGTACAGCA 20
          ||||||||||||||||||||
Db      1 CCAACAGGCACAGTACAGCA 20
```

RESULT 2
AAL33140/c
ID AAL33140 standard; DNA; 51 BP.
XX
AC AAL33140;
XX
DT 24-JAN-2002 (first entry)
XX
DE Human SNP oligonucleotide #6348.
XX
KW Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
KW neuroprotective; antimicrobial; gene therapy; vaccine; amylase; cancer;
KW amyloid protein; angiopoietin; apoptosis related protein; cadherin;
KW cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
KW complement related protein; cytochrome; kinesin; cytokine; interferon;
KW interleukin; G-protein coupled receptor; thioesterase; inflammation;
KW multifactorial disease; autoimmune disease; infection;
KW nervous system disease; ss.
XX
OS Homo sapiens.
XX
PN WO200147944-A2.
XX
PD 05-JUL-2001.
XX
PF 28-DEC-2000; 2000WO-US35498.
XX
PR 28-DEC-1999; 99US-0173419.
PR 27-DEC-2000; 2000US-0173419.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shimkets RA, Leach M;
XX
DR WPI; 2001-465210/50.
XX
PT Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
PT oncogenes and histones, useful for diagnosing and treating, e.g.
PT cancer, autoimmune diseases and infections -
XX
PS Claim 1; Page 3201; 4143pp; English.
XX
CC The present invention relates to oligonucleotides encoding polymorphic
CC variants of proteins related to amylases, amyloid proteins, angiopoietin,
CC apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,
CC histones, kinases, colony stimulating factors, complement related
CC proteins, cytochromes, kinesins, cytokines, interferons, interleukins,
CC G-protein coupled receptors and thioesterases. The present sequence is
CC one such oligonucleotide. The oligonucleotides and the peptides encoded
CC by them may be used in the prevention, diagnosis and treatment of
CC diseases associated with inappropriate expression of the proteins listed
CC above. Disorders that may be prevented, diagnosed and/or treated include
CC multifactorial diseases with a genetic component, such as autoimmune
CC diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,
CC systemic lupus erythromatosus and Grave's disease), inflammation, cancer
CC (e.g. cancers of the bladder, brain, breast, colon and kidney,
CC leukaemia), diseases of the nervous system and an infection of pathogenic
CC organisms.
XX
SQ Sequence 51 BP; 12 A; 11 C; 10 G; 18 T; 0 other;

Query Match 84.0%; Score 16.8; DB 22; Length 51;
Best Local Similarity 90.0%; Pred. No. 1.2e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy      1 CCAACAGGCACAGTACAGCA 20
          ||  ||||||||||||||||
Db     46 CCTGCAGGCACAGTACAGCA 27
```

RESULT 3
AAL33141/c
ID AAL33141 standard; DNA; 51 BP.
XX
AC AAL33141;
XX
DT 24-JAN-2002 (first entry)
XX
DE Human SNP oligonucleotide #6349.
XX
KW Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
KW neuroprotective; antimicrobial; gene therapy; vaccine; amylase; cancer;
KW amyloid protein; angiopoietin; apoptosis related protein; cadherin;
KW cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
KW complement related protein; cytochrome; kinesin; cytokine; interferon;
KW interleukin; G-protein coupled receptor; thioesterase; inflammation;
KW multifactorial disease; autoimmune disease; infection;
KW nervous system disease; ss.
XX
OS Homo sapiens.
XX
PN WO200147944-A2.
XX
PD 05-JUL-2001.
XX
PF 28-DEC-2000; 2000WO-US35498.
XX
PR 28-DEC-1999; 99US-0173419.
PR 27-DEC-2000; 2000US-0173419.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shimkets RA, Leach M;
XX
DR WPI; 2001-465210/50.
XX
PT Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
PT oncogenes and histones, useful for diagnosing and treating, e.g.
PT cancer, autoimmune diseases and infections -
XX
PS Claim 1; Page 3201; 4143pp; English.
XX
CC The present invention relates to oligonucleotides encoding polymorphic
CC variants of proteins related to amylases, amyloid proteins, angiopoietin,
CC apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,
CC histones, kinases, colony stimulating factors, complement related
CC proteins, cytochromes, kinesins, cytokines, interferons, interleukins,
CC G-protein coupled receptors and thioesterases. The present sequence is

CC one such oligonucleotide. The oligonucleotides and the peptides encoded
CC by them may be used in the prevention, diagnosis and treatment of
CC diseases associated with inappropriate expression of the proteins listed
CC above. Disorders that may be prevented, diagnosed and/or treated include
CC multifactorial diseases with a genetic component, such as autoimmune
CC diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,
CC systemic lupus erythromatosus and Gravé's disease), inflammation, cancer
CC (e.g. cancers of the bladder, brain, breast, colon and kidney,
CC leukaemia), diseases of the nervous system and an infection of pathogenic
CC organisms.
XX
SQ Sequence 51 BP; 12 A; 12 C; 12 G; 15 T; 0 other;

Query Match 79.0%; Score 15.8; DB 22; Length 51;
Best Local Similarity 89.5%; Pred. No. 3.4e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CCAACAGGCACAGTACAGC 19
|| |||||||||||||||
Db 45 CCTGCAGGCACAGTACAGC 27

RESULT 4
AAL33142/c
ID AAL33142 standard; DNA; 51 BP.
XX
AC AAL33142;
XX
DT 24-JAN-2002 (first entry)
XX
DE Human SNP oligonucleotide #6350.
XX
KW Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
KW neuroprotective; antimicrobial; gene therapy; vaccine; amylase; cancer;
KW amyloid protein; angiopoietin; apoptosis related protein; cadherin;
KW cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
KW complement related protein; cytochrome; kinesin; cytokine; interferon;
KW interleukin; G-protein coupled receptor; thioesterase; inflammation;
KW multifactorial disease; autoimmune disease; infection;
KW nervous system disease; ss.
XX
OS Homo sapiens.
XX
PN WO200147944-A2.
XX
PD 05-JUL-2001.
XX
PF 28-DEC-2000; 2000WO-US35498.
XX
PR 28-DEC-1999; 99US-0173419.
PR 27-DEC-2000; 2000US-0173419.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shimkets RA, Leach M;
XX
DR WPI; 2001-465210/50.
XX
PT Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
PT oncogenes and histones, useful for diagnosing and treating, e.g.
PT cancer, autoimmune diseases and infections -
XX
PS Claim 1; Page 3201; 4143pp; English.
XX
CC The present invention relates to oligonucleotides encoding polymorphic
CC variants of proteins related to amylases, amyloid proteins, angiopoietin,
CC apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,
CC histones, kinases, colony stimulating factors, complement related
CC proteins, cytochromes, kinesins, cytokines, interferons, interleukins,
CC G-protein coupled receptors and thioesterases. The present sequence is
CC one such oligonucleotide. The oligonucleotides and the peptides encoded
CC by them may be used in the prevention, diagnosis and treatment of
CC diseases associated with inappropriate expression of the proteins listed
CC above. Disorders that may be prevented, diagnosed and/or treated include
CC multifactorial diseases with a genetic component, such as autoimmune
CC diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,
CC systemic lupus erythromatosus and Grave's disease), inflammation, cancer
CC (e.g. cancers of the bladder, brain, breast, colon and kidney,
CC leukaemia), diseases of the nervous system and an infection of pathogenic
CC organisms.
XX
SQ Sequence 51 BP; 10 A; 12 C; 15 G; 14 T; 0 other;

Query Match 76.0%; Score 15.2; DB 22; Length 51;
Best Local Similarity 85.0%; Pred. No. 6.6e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CCAACAGGCACAGTACAGCA 20
|| ||| ||||||||||||
Db 33 CCTGCAGACACAGTACAGCA 14

RESULT 5
ABN80383/c
ID ABN80383 standard; DNA; 24 BP.
XX
AC ABN80383;
XX
DT 16-JUL-2002 (first entry)
XX
DE Oligonucleotide primer #7.
XX
KW Amplification; chimeric oligonucleotide primer; disease diagnosis;
KW polymerase chain reaction; PCR; genetic engineering; blood; urine;
KW plant tissue; animal tissue; assay; soil; food; microorganism; ss.
XX
OS Escherichia coli.
XX
PN WO200216639-A1.
XX
PD 28-FEB-2002.
XX
PF 21-AUG-2001; 2001WO-JP07139.
XX
PR 23-AUG-2000; 2000JP-0251981.
PR 19-SEP-2000; 2000JP-0284419.
PR 22-SEP-2000; 2000JP-0288750.
PR 03-APR-2001; 2001JP-0104191.
XX
PA (TAKI) TAKARA SHUZO CO LTD.
XX
PI Sagawa H, Uemori T, Mukai H, Yamamoto J, Tomono J, Kobayashi E;
PI Enoki T, Asada K, Kato I;
XX
DR WPI; 2002-351653/38.
XX
PT Amplifying a target nucleic acid in sample, useful in e.g. clinical
PT applications, genetic engineering and for assaying blood, urine, plant
PT and animal tissues and environmental materials like soil and food -
XX
PS Examples; Page 259; 332pp; Japanese.
XX
CC The invention relates to the amplification of a nucleic acid. This
CC comprises using a nucleic acid as template, deoxypolynucleotide
CC 3-phosphate, a chimeric oligonucleotide primer with a ribonucleotide
CC provided at the 3'-terminus or in the 3'-terminal side, DNA polymerase
CC with a chain-transfer activity, an RNaseH or endonuclease, and incubating
CC the mixture to give a reaction product. The method is useful for
CC amplifying a target nucleic acid in a sample, which is useful in e.g.
CC clinical applications including disease diagnosis, genetic engineering,
CC in assaying blood, urine, plant and animal tissues, environmental
CC materials like soil and food and identification of microorganisms. The
CC method of the invention, known as an isothermal and chimeric
CC primer-initiated amplification of nucleic acids (ICAN) method, is highly
CC sensitive and specific. Sequences given in records ABN80338-ABN80532
CC represent target nucleic acids, chimeric oligonucleotides and

```
CC   oligonucleotide primers and probes of the invention. Chimeric
CC   oligonucleotides are DNA/RNA hybrids.
XX
SQ   Sequence 24 BP; 5 A; 4 C; 7 G; 8 T; 0 other;

  Query Match             74.0%;  Score 14.8;  DB 24;  Length 24;
  Best Local Similarity   88.9%;  Pred. No. 9.4e+02;
  Matches   16;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       3 AACAGGCACAGTACAGCA 20
           ||||||||||||||  ||
Db      21 AACAGGCACAGTACCCCA 4


RESULT 6
AAH42148/c
ID   AAH42148 standard; DNA; 52 BP.
XX
AC   AAH42148;
XX
DT   17-SEP-2001  (first entry)
XX
DE   HRE element from the endothelin (ET-1) gene.
XX
KW   Expression vector; silencer element; inducible element;
KW   silencer-inducible region; gene therapy; cardiac disease;
KW   immunodeficiency; allergy; anemia; thalessemia; autoimmune disease;
KW   shock; hemophilia; inflammation; stress; ischemia; hypoxic condition;
KW   carcinoma; leukemia; Hodgkin disease; Kaposi sarcoma;
KW   hypoxia response enhancer; HRE; endothelin gene; ET-1 gene; ss.
XX
OS   Unidentified.
XX
PN   WO200148187-A2.
XX
PD   05-JUL-2001.
XX
PF   15-DEC-2000; 2000WO-US33269.
XX
PR   23-DEC-1999;   99US-0171597.
PR   28-NOV-2000; 2000US-0723326.
XX
PA   (UYMI-) UNIV MIAMI.
XX
PI   Webster KA;
XX
DR   WPI; 2001-441715/47.
XX
PT   Novel isolated expression vector useful therapeutically, comprisessa
PT   silencer elements and conditionally inducible elements to form
PT   silencer-inducible region, and a promoter in operative linkage with the
PT   region -
XX
PS   Disclosure; Page 26; 49pp; English.
XX
CC   The specification describes an expression vector. The vector comprises
CC   silencer elements and conditionally inducible elements to form a
CC   silencer-inducible region (IR), and a promoter in operative linkage
CC   with IR, where the promoter is regulated by IR, and upstream of the
CC   expressed region. The vector is useful diagnostically, therapeutically,
CC   prophylactically to make models of human disease. It is useful in gene
CC   therapy, production of recombinant biologicals, genetic diagnosis, drug
CC   screening, and genetic research (e.g., genomics, proteomics, in vivo
CC   and in vitro models of human disease). It is useful for treating cardiac
CC   disease (by reduction or prevention of ischemic damage, inhibition of
CC   restenosis, neutralization of other pathological effects of heart or
CC   vascular disease, or diagnosis of hypoxia), acquired or inherited
CC   immunodeficiency, allergy, anemia, thalessemia, autoimmune disease,
CC   hemolytic or septic shock, hemophilia, inflammation and other stress
CC   conditions, ischemia and other hypoxic conditions, carcinoma, leukemia,
CC   Hodgkin disease, non-Hodgkin lymphoma and Kaposi sarcoma. It is also
CC   useful for suppressing or eliminating infectious agents, autoimmune cells
CC   and cancerous cells, and for preventing an infection or disease in a
CC   patient. The present sequence represents an oligonucleotide containing
CC   the hypoxia response enhancer (HRE) element from the endothelin (ET-1)
CC   gene. It is used to produce vectors of the invention.
XX
SQ   Sequence 52 BP; 12 A; 14 C; 11 G; 15 T; 0 other;

  Query Match             71.0%;  Score 14.2;  DB 22;  Length 52;
  Best Local Similarity   84.2%;  Pred. No. 2e+03;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 CCAACAGGCACAGTACAGC 19
           ||||||||||  || ||||
Db      34 CCAACAGGCAACGTGCAGC 16


RESULT 7
ABN52993/c
ID   ABN52993 standard; DNA; 65 BP.
XX
AC   ABN52993;
XX
DT   15-JUL-2002  (first entry)
XX
DE   Mouse spliced transcript detection oligonucleotide SEQ ID NO:25741.
XX
KW   Human; mouse; rat; splice transcript; detection; RNA transcript;
KW   splice variant; transcriptome; oligonucleotide library; ss.
XX
OS   Mus musculus.
XX
PN   WO200210449-A2.
XX
PD   07-FEB-2002.
XX
PF   20-JUL-2001; 2001WO-IB01903.
XX
PR   28-JUL-2000; 2000US-221607P.
PR   02-MAY-2001; 2001US-287724P.
XX
PA   (COMP-) COMPUGEN INC.
XX
PI   Shoshan A,  Wasserman A,  Mintz E,  Mintz L,  Faigler S;
XX
DR   WPI; 2002-257383/30.
XX
PT   New oligonucleotide libraries comprising oligonucleotides which
PT   selectively hybridize to mRNAs transcribed from a transcription unit of
PT   a genome, useful for detecting tissue-, pathology-, and
PT   developmental-specific genes -
XX
PS   Example 1; SEQ ID 25741; 47pp; English.
XX
CC   The present invention describes oligonucleotide libraries for detecting
CC   messenger RNAs that populate a (sub-)transcriptome, where the
CC   (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC   transcription units that populate a genome. The library comprises
CC   several oligonucleotides, each capable of hybridising selectively to a
CC   set of messenger RNAs transcribed from a given transcription unit of
CC   the genome, which encodes one or more messenger RNA splice variants.
CC   The oligonucleotide libraries are useful for detecting mRNAs from a
CC   biological sample, in expression profiling studies, in qualitatively or
CC   quantitatively characterising the corresponding transcriptome, and in
CC   detecting RNA transcripts and splice variants of human or animal
CC   transcriptomes. The libraries may also be used as specialised mini
CC   libraries to detect transcripts of a sub-transcriptome under a
CC   particular biological or pathological state, and so allowing the
CC   detection of tissue- and pathology-specific genes such as those genes
CC   only expressed in specific tissue under a specific pathological
CC   condition; to detect developmental specific genes; and to detect RNA
CC   transcripts and splice variants of a transcriptome of a patient suffering
CC   from a particular disorder. ABN27253 to ABN59589 represent
CC   oligonucleotide sequences from rats, humans and mice, which are used in
CC   the exemplification of the present invention.
```

```
CC   N.B. The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 65 BP; 10 A; 16 C; 18 G; 21 T; 0 other;

  Query Match            71.0%;  Score 14.2;  DB 24;  Length 65;
  Best Local Similarity  84.2%;  Pred. No. 2e+03;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 CCAACAGGCACAGTACAGC 19
          |||| |||||||| | |||
Db     54 CCAAAAGGCACAGCAGAGC 36
```

```
RESULT 8
ABL45353
ID   ABL45353 standard; DNA; 20 BP.
XX
AC   ABL45353;
XX
DT   11-APR-2002  (first entry)
XX
DE   Human chromosome 21q22.1 PCR primer SEQ ID NO:2397.
XX
KW   Human; chromosome 1p36-35; chromosome 21q22.1; genetic analysis;
KW   genome; PCR primer; ss.
XX
OS   Homo sapiens.
XX
PN   JP2001321190-A.
XX
PD   20-NOV-2001.
XX
PF   12-MAR-2001; 2001JP-0068285.
XX
PR   10-MAR-2000; 2000JP-0066716.
XX
PA   (RIKA ) RIKAGAKU KENKYUSHO.
PA   (GENO-) GENOTEX YG.
XX
DR   WPI; 2002-144136/19.
XX
PT   Arraying genome clones  -
XX
PS   Claim 6; Page 52; 528pp; Japanese.
XX
CC   The present invention describes a method of arraying genome clones. The
CC   method comprises: (a) clones of the genomic libraries contained in
CC   multiwell plates numbered for discrimination are mixed in each of the
CC   multiwell plates; (b) a primer designed based on the chromosome marker
CC   sequence is added to the mixture to carry out an amplification reaction;
CC   (c) a signal corresponding to the marker is detected from the resultant
CC   amplified product to specify the discrimination Nos. of the multiwell
CC   plates containing the clones having said marker sequence; (d) the order
CC   of the markers is changed so that the same discrimination Nos. succeed to
CC   the maximum in the specified discrimination Nos. to array the multiwell
CC   plates; (e) the clones in the multiwell plates of the specified
CC   discrimination Nos. are mixed respectively in each wells of longitudinal
CC   and lateral directions; (f) the mixed clones are cultured and the
CC   resultant cultures are amplified by using the above primer; (g) signals
CC   are detected from the amplified products; (h) the clones in the multiwell
CC   plates are specified from the detected result; and (i) the clones are
CC   reconstituted as the positions on the chromosome and arrayed. The
CC   microarray is useful for gene analysis. ABL42957 to ABL45322 represent
CC   PCR primers for human chromosome 1p36-35 DNA, and ABL45323 to ABL45634
CC   represent PCR primers for human chromosome 21q22.1, which are
CC   specifically claimed for use in the present invention.
XX
SQ   Sequence 20 BP; 5 A; 4 C; 8 G; 3 T; 0 other;

  Query Match            70.0%;  Score 14;  DB 24;  Length 20;
  Best Local Similarity  100.0%;  Pred. No. 2.2e+03;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      5 CAGGCACAGTACAG 18
          ||||||||||||||
Db      3 CAGGCACAGTACAG 16
```

```
RESULT 9
AAF74881
ID   AAF74881 standard; DNA; 21 BP.
XX
AC   AAF74881;
XX
DT   22-MAY-2001  (first entry)
XX
DE   Chicken neuropilin-1 mutagenesis PCR primer SEQ ID NO:4.
XX
KW   Chicken; neuropilin-1; secreted semaphorin receptor; collapse; motility;
KW   axon growth core; axon generation; mutagenesis; PCR primer; ss.
XX
OS   Gallus gallus.
XX
PN   WO200118173-A2.
XX
PD   15-MAR-2001.
XX
PF   08-SEP-2000; 2000WO-US24635.
XX
PR   10-SEP-1999;   99US-0153309.
PR   16-DEC-1999;   99US-0171176.
XX
PA   (UYPE-) UNIV PENNSYLVANIA.
XX
PI   Raper JA,  Renzi MJ;
XX
DR   WPI; 2001-235194/24.
XX
PT   New dominant negative neuropilin-1 receptor, useful for modulating or
PT   inactivating activity of selected secreted semaphorins and inhibiting
PT   or preventing collapse or motility of axon growth cone  -
XX
PS   Example 1; Page 27; 68pp; English.
XX
CC   The present invention describes an isolated DNA encoding a dominant
CC   negative receptor (I), where the DNA comprises a nucleic acid sequence
CC   encoding a neuropilin-1 which has semaphorin receptor specific
CC   antigenicity or immunogenicity, including homologues, modifications,
CC   derivatives and active fragments. Also described is a protein comprising
CC   a dominant negative receptor, which has receptor-specific antigenicity
CC   or immunogenicity for semaphorin 3A and semaphorin 3C, but not for
CC   semaphorin 3F. (I) Is useful for modulating or inactivating the activity
CC   of selected secreted semaphorins, for inhibiting or preventing the
CC   collapse or motility of an axon growth cone, where motility or collapse
CC   is mediated by a secreted semaphorin, for modulating overgrowth or
CC   premature entry of axons to their targets in vivo, and for enhancing
CC   axon generation or regeneration by blocking secreted semaphorin
CC   binding, by adding or overexpressing (I). The axon growth occurs in a
CC   developing or regenerating neurological system. The present sequence
CC   represents a PCR primer which is used in an example from the present
CC   invention.
XX
SQ   Sequence 21 BP; 8 A; 6 C; 6 G; 1 T; 0 other;

  Query Match            70.0%;  Score 14;  DB 22;  Length 21;
  Best Local Similarity  100.0%;  Pred. No. 2.2e+03;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      5 CAGGCACAGTACAG 18
          ||||||||||||||
Db      2 CAGGCACAGTACAG 15
```

```
RESULT 10
```

```
AAC91777/c
ID   AAC91777 standard; DNA; 21 BP.
XX
AC   AAC91777;
XX
DT   27-MAR-2001  (first entry)
XX
DE   Escherichia coli O157 verotoxin 2 (VT2) PCR primer, SEQ ID NO:1.
XX
KW   Verotoxin 2; VT2; strain O157; nucleic acid-immobilised substrate;
KW   nucleic acid array; DNA chip; carbodiimide group; isocyanate group;
KW   PCR primer; ss.
XX
OS   Escherichia coli.
XX
PN   EP1063301-A2.
XX
PD   27-DEC-2000.
XX
PF   20-JUN-2000; 2000EP-0305190.
XX
PR   21-JUN-1999;   99JP-0173966.
XX
PA   (NISN ) NISSHINBO IND INC.
XX
PI   Suzuki O,  Ichihara T,  Shiohata N,  Matsumura Y;
XX
DR   WPI; 2001-082770/10.
XX
PT   Substrate carrying immobilized nucleic acid molecule, useful as DNA
PT   array, comprises a carrier comprising a base material and a compound
PT   having carbodiimide or isocyanate group, and nucleic acids  -
XX
PS   Example 1; Page 19; 21pp; English.
XX
CC   The invention relates to a novel nucleic acid-immobilised substrate,
CC   which may be used as a DNA array or DNA chip. The nucleic acid-
CC   immobilised substrate comprises a carrier, consisting of a base material
CC   to which a compound with a carbodiimide or an isocyanate group is
CC   attached, and a plurality of nucleic acids which are securely
CC   immobilised on the carrier in the form of dots via the carbodiimide or
CC   isocyanate moieties. The nucleic acids are stably immobilised on the
CC   substrate without any limitations concerning the number of chains or
CC   length of the nucleic acids, and thus various kinds of nucleic acids can
CC   simultaneously be handled on the same base material. Since the nucleic
CC   acids are securely bound to the carrier through covalent bonds, nucleic
CC   acid-immobilised substrates are useful as DNA chips, having excellent
CC   reproducibility and quantification characteristics. The present sequence
CC   represents an Escherichia coli O157 verotoxin 2 (VT2) PCR primer used in
CC   an exemplification of the invention.
XX
SQ   Sequence 21 BP; 5 A; 3 C; 6 G; 7 T; 0 other;

  Query Match             70.0%;  Score 14;  DB 22;  Length 21;
  Best Local Similarity   100.0%;  Pred. No. 2.2e+03;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       3 AACAGGCACAGTAC 16
           ||||||||||||||
Db      20 AACAGGCACAGTAC 7


RESULT 11
AAF98589/c
ID   AAF98589 standard; DNA; 22 BP.
XX
AC   AAF98589;
XX
DT   02-JUL-2001  (first entry)
XX
DE   Human kinase marker 25 reverse primer.
XX
KW   Human; ovarian cancer; identification; detection; characterisation;
KW   tumour; kinase; marker; cytostatic; antisense gene therapy; probe;
KW   primer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200118542-A2.
XX
PD   15-MAR-2001.
XX
PF   01-SEP-2000; 2000WO-US24199.
XX
PR   03-SEP-1999;   99US-0152547.
PR   16-MAR-2000; 2000US-0190347.
PR   21-MAR-2000; 2000US-0191321.
PR   31-MAY-2000; 2000US-0208382.
PR   20-JUL-2000; 2000US-0220467.
XX
PA   (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI   Lee J,  Thompsho P,  Lillie J;
XX
DR   WPI; 2001-211428/21.
XX
PT   Detection, assessment, prevention and therapy of ovarian cancer,
PT   comprises detecting changes in the expression of a variety of markers -
XX
PS   Disclosure; Page 102; 1198pp; English.
XX
CC   The present invention describes a method for assessing whether a patient
CC   is afflicted with ovarian cancer by comparing: (1) the expression of a
CC   marker (I) (see AAF98594 to AAF98730), in a patient sample; and (2) the
CC   normal level of expression of (I) in a control non-ovarian cancer
CC   sample, where a significant difference between the level of expression
CC   in (a) and (b) is an indication that the patient is afflicted with
CC   ovarian cancer. (I) have cytostatic activities and can be used in
CC   antisense gene therapy. The method, compositions and kits from the
CC   present invention can be used for: (1) assessing and treating ovarian
CC   cancer; (2) making isolated hybridoma, which produces an antibody useful
CC   for ovarian cancer assessment; and (3) inhibiting ovarian cancer in a
CC   patient. AAF98573 to AAF98593 represent human kinase marker primers and
CC   probes which are used in the exemplification of the present invention.
XX
SQ   Sequence 22 BP; 2 A; 7 C; 5 G; 8 T; 0 other;

  Query Match             69.0%;  Score 13.8;  DB 22;  Length 22;
  Best Local Similarity   88.2%;  Pred. No. 2.8e+03;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       3 AACAGGCACAGTACAGC 19
           ||| |||| ||||||||
Db      18 AACCGGCAGAGTACAGC 2


RESULT 12
AAA04344/c
ID   AAA04344 standard; DNA; 29 BP.
XX
AC   AAA04344;
XX
DT   22-MAY-2000  (first entry)
XX
DE   Polymorphic fragment of hypertension associated gene GSY1.
XX
KW   Polymorphism; hypertension; agammaglobulinemia; diabetes insipidus;
KW   Lesch-Nyhan syndrome; muscular dystrophy; Wiskott-Aldrich syndrome;
KW   Fabrys disease; familial hypercholesterolemia; hereditary spherocytosis;
KW   polycystic kidney disease; von Willebrands disease;  forensic; human;
KW   tuberous sclerosis; hereditary hemorrhagica telangiectasia;
KW   familial colonic polyposis; osteogenesis imperfecta; porphyria;
KW   Ehlers-Danlos syndrome; ss.
XX
OS   Homo sapiens.
XX
```

PN EP955382-A2.
XX
PD 10-NOV-1999.
XX
PF 07-MAY-1999; 99EP-0250150.
XX
PR 07-MAY-1998; 98US-0084641.
PR 03-MAY-1999; 99US-0304232.
XX
PA (AFFY-) AFFYMETRIX INC.
PA (UYCA-) UNIV CASE WESTERN RESERVE.
XX
PI Fan JB, Chakravarti A, Haluska MK;
XX
DR WPI; 2000-107928/10.
XX
PT Novel nucleic acids containing polymorphisms used in the diagnosis of
PT hypertension -
XX
PS Claim 1; Page 33; 53pp; English.
XX
CC The invention provides polymorphic fragments of genes associated with
CC hypertension. The nucleic acids including the polymorphic sites can be
CC used as probes or primers for expressing variant proteins. Detection of
CC the polymorphisms is useful in designing prophylactic and therapeutic
CC regimes customized to underlying abnormalities. The polymorphisms can be
CC used for association studies for hypertension, and in hypertension
CC diagnostic assays. Where the polymorphisms have strong correlation with
CC hypertension, within a gene, they are likely to have a causative role in
CC hypertension. This information can be used to find the precise role of a
CC polymorphism in the disease, and this can be used to identify potential
CC drugs which combat the disease. The polymorphisms can be tested for
CC association with other diseases e.g. agammaglobulinemia, diabetes
CC insipidus, Lesch-Nyhan syndrome, muscular dystrophy, Wiskott-Aldrich
CC syndrome, Fabrys disease, familial hypercholesterolemia, polycystic
CC kidney disease, hereditary spherocytosis, von Willebrands disease,
CC tuberous sclerosis, hereditary hemorrhagica telangiectasia, familial
CC colonic polyposis, Ehlers-Danlos syndrome, osteogenesis imperfecta, and
CC acute intermittent porphyria. The polymorphic forms can also be used in
CC forensics to identify individuals.
XX
SQ Sequence 29 BP; 2 A; 6 C; 9 G; 11 T; 1 other;

Query Match 69.0%; Score 13.8; DB 21; Length 29;
Best Local Similarity 78.9%; Pred. No. 2.8e+03;
Matches 15; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CCAACAGGCACAGTACAGC 19
||| :|||||| ||| |||
Db 19 CCACMAGGCACGGTAAAGC 1

RESULT 13
AAH38884/c
ID AAH38884 standard; DNA; 51 BP.
XX
AC AAH38884;
XX
DT 14-AUG-2001 (first entry)
XX
DE Human SNP flanking oligonucleotide SEQ ID 1680.
XX
KW Single nucleotide polymorphism; SNP; single nucleotide primer extension;
KW SNPE; genotyping; agammaglobulinaemia; diabetes insipidus; cancer;
KW Lesch-Nyhan syndrome; muscular dystrophy; familial hypercholesterolaemia;
KW polycystic kidney disease; osteogenesis imperfecta; autoimmune disease;
KW acute intermittent porphyria; rheumatoid arthritis; multiple sclerosis;
KW inflammation; forensic investigation; paternity analysis; ds.
XX
OS Homo sapiens.
XX
PN WO200129262-A2.
XX
PD 26-APR-2001.
XX
PF 13-OCT-2000; 2000WO-US28436.
XX
PR 15-OCT-1999; 99US-0160096.
XX
PA (ORCH-) ORCHID BIOSCIENCES INC.
XX
PI Picoult-Newburg L, Pohl M;
XX
DR WPI; 2001-290930/30.
XX
PT New genotyping oligonucleotide, useful for detecting the presence,
PT absence or identity of single polynucleotide polymorphism in a nucleic
PT acid sample -
XX
PS Claim 1; Page 58; 83pp; English.
XX
CC Sequences AAH37205 - AAH40944 represent PCR primers, single nucleotide
CC primer extension (SNPE) primers, and the sequences of regions flanking
CC sites of single nucleotide polymorphisms SNPs. The present invention
CC includes kits for determining the presence or absence of a SNP, using the
CC oligonucleotides of the invention. The PCR primers are used to amplify a
CC SNP flanking sequence, the SNPE primer is used as a genotyping primer.
CC The oligonucleotides are useful for genotyping a nucleic acid sample by
CC performing a single-nucleotide primer extension reaction. The
CC oligonucleotides are useful for determining the presence, absence or
CC identity of a SNP and for genotyping nucleic acid samples, for e.g. to
CC assess by association analysis the genotype of an individual or group of
CC individuals, having a pathological phenotypic trait suspected of being
CC caused by one or more SNPs. Phenotypic traits include diseases e.g.
CC agammaglobulinaemia, diabetes insipidus, Lesch-Nyhan syndrome, muscular
CC dystrophy, familial hypercholesterolaemia, polycystic kidney disease,
CC osteogenesis imperfecta and acute intermittent porphyria. Phenotypic
CC traits also include symptoms of or susceptibility to multifactorial
CC disease of which a component is or may be genetic such as autoimmune
CC diseases, including, rheumatoid arthritis, multiple sclerosis,
CC inflammation, cancer, nervous system diseases and infection by pathogenic
CC microorganism. The method is also useful in forensic investigations and
CC paternity analysis. The present sequence represents a fragment of human
CC DNA flanking the site of a single nucleotide polymorphism.
XX
SQ Sequence 51 BP; 15 A; 5 C; 18 G; 13 T; 0 other;

Query Match 69.0%; Score 13.8; DB 22; Length 51;
Best Local Similarity 88.2%; Pred. No. 3e+03;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CCAACAGGCACAGTACA 17
||||||||||| | ||
Db 36 CCAACAGGCACATTCCA 20

RESULT 14
ABN34175/c
ID ABN34175 standard; DNA; 60 BP.
XX
AC ABN34175;
XX
DT 15-JUL-2002 (first entry)
XX
DE Human spliced transcript detection oligonucleotide SEQ ID NO:6923.
XX
KW Human; mouse; rat; splice transcript; detection; RNA transcript;
KW splice variant; transcriptome; oligonucleotide library; ss.
XX
OS Homo sapiens.
XX
PN WO200210449-A2.
XX
PD 07-FEB-2002.
XX
PF 20-JUL-2001; 2001WO-IB01903.

```
XX
PR   28-JUL-2000; 2000US-221607P.
PR   02-MAY-2001; 2001US-287724P.
XX
PA   (COMP-) COMPUGEN INC.
XX
PI   Shoshan A,  Wasserman A,  Mintz E,  Mintz L,  Faigler S;
XX
DR   WPI; 2002-257383/30.
XX
PT   New oligonucleotide libraries comprising oligonucleotides which
PT   selectively hybridize to mRNAs transcribed from a transcription unit of
PT   a genome, useful for detecting tissue-, pathology-, and
PT   developmental-specific genes  -
XX
PS   Example 1; SEQ ID 6923; 47pp; English.
XX
CC   The present invention describes oligonucleotide libraries for detecting
CC   messenger RNAs that populate a (sub-)transcriptome, where the
CC   (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC   transcription units that populate a genome. The library comprises
CC   several oligonucleotides, each capable of hybridising selectively to a
CC   set of messenger RNAs transcribed from a given transcription unit of
CC   the genome, which encodes one or more messenger RNA splice variants.
CC   The oligonucleotide libraries are useful for detecting mRNAs from a
CC   biological sample, in expression profiling studies, in qualitatively or
CC   quantitatively characterising the corresponding transcriptome, and in
CC   detecting RNA transcripts and splice variants of human or animal
CC   transcriptomes. The libraries may also be used as specialised mini
CC   libraries to detect transcripts of a sub-transcriptome under a
CC   particular biological or pathological state, and so allowing the
CC   detection of tissue- and pathology-specific genes such as those genes
CC   only expressed in specific tissue under a specific pathological
CC   condition; to detect developmental specific genes; and to detect RNA
CC   transcripts and splice variants of a transcriptome of a patient suffering
CC   from a particular disorder. ABN27253 to ABN59589 represent
CC   oligonucleotide sequences from rats, humans and mice, which are used in
CC   the exemplification of the present invention.
CC   N.B. The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 60 BP; 10 A; 16 C; 14 G; 20 T; 0 other;

  Query Match             69.0%;  Score 13.8;  DB 24;  Length 60;
  Best Local Similarity   88.2%;  Pred. No. 3.1e+03;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 CCAACAGGCACAGTACA 17
           | ||||||  ||||||||
Db      54 CTAACAGGAACAGTACA 38


RESULT 15
ABN41646
ID   ABN41646 standard; DNA; 60 BP.
XX
AC   ABN41646;
XX
DT   15-JUL-2002  (first entry)
XX
DE   Human spliced transcript detection oligonucleotide SEQ ID NO:14394.
XX
KW   Human; mouse; rat; splice transcript; detection; RNA transcript;
KW   splice variant; transcriptome; oligonucleotide library; ss.
XX
OS   Homo sapiens.
XX
PN   WO200210449-A2.
XX
PD   07-FEB-2002.
XX
PF   20-JUL-2001; 2001WO-IB01903.
```

```
XX
PR   28-JUL-2000; 2000US-221607P.
PR   02-MAY-2001; 2001US-287724P.
XX
PA   (COMP-) COMPUGEN INC.
XX
PI   Shoshan A,  Wasserman A,  Mintz E,  Mintz L,  Faigler S;
XX
DR   WPI; 2002-257383/30.
XX
PT   New oligonucleotide libraries comprising oligonucleotides which
PT   selectively hybridize to mRNAs transcribed from a transcription unit of
PT   a genome, useful for detecting tissue-, pathology-, and
PT   developmental-specific genes  -
XX
PS   Example 1; SEQ ID 14394; 47pp; English.
XX
CC   The present invention describes oligonucleotide libraries for detecting
CC   messenger RNAs that populate a (sub-)transcriptome, where the
CC   (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC   transcription units that populate a genome. The library comprises
CC   several oligonucleotides, each capable of hybridising selectively to a
CC   set of messenger RNAs transcribed from a given transcription unit of
CC   the genome, which encodes one or more messenger RNA splice variants.
CC   The oligonucleotide libraries are useful for detecting mRNAs from a
CC   biological sample, in expression profiling studies, in qualitatively or
CC   quantitatively characterising the corresponding transcriptome, and in
CC   detecting RNA transcripts and splice variants of human or animal
CC   transcriptomes. The libraries may also be used as specialised mini
CC   libraries to detect transcripts of a sub-transcriptome under a
CC   particular biological or pathological state, and so allowing the
CC   detection of tissue- and pathology-specific genes such as those genes
CC   only expressed in specific tissue under a specific pathological
CC   condition; to detect developmental specific genes; and to detect RNA
CC   transcripts and splice variants of a transcriptome of a patient suffering
CC   from a particular disorder. ABN27253 to ABN59589 represent
CC   oligonucleotide sequences from rats, humans and mice, which are used in
CC   the exemplification of the present invention.
CC   N.B. The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 60 BP; 18 A; 13 C; 13 G; 16 T; 0 other;

  Query Match             69.0%;  Score 13.8;  DB 24;  Length 60;
  Best Local Similarity   88.2%;  Pred. No. 3.1e+03;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       4 ACAGGCACAGTACAGCA 20
           |||||||  ||||||||
Db      32 ACAGGCAATGTACAGCA 48


Search completed: November 23, 2002, 07:04:00
Job time : 97.1 secs
```

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 26, 2002, 04:16:10 ; Search time 798.5 Seconds
(without alignments)
405.648 Million cell updates/sec

Title: US-09-296-264-28
Perfect score: 20
Sequence: 1 ccaacaggcacagtacagca 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 357874

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	14.8	74.0	100	10	AW454338	AW454338 zeh11308
c 2	14.2	71.0	69	9	AL632098	AL632098 AL632098
c 3	14.2	71.0	73	17	AZ662265	AZ662265 1M0541F15
c 4	14.2	71.0	78	10	AW477033	AW477033 ga40h10.y
c 5	14	70.0	71	10	AV914039	AV914039 AV914039
c 6	13.8	69.0	51	17	AZ779007	AZ779007 2M0014H16
c 7	13.8	69.0	86	14	BQ760503	BQ760503 EBro02_SQ
8	13.8	69.0	91	14	D18618	D18618 MUSGS01679
9	13.8	69.0	94	17	CNS02A9J	AL188272 Tetraodon
c 10	13.6	68.0	87	14	BQ807426	BQ807426 NISC_kk03
11	13.6	68.0	90	14	F32010	F32010 HSPD23880 H
c 12	13.6	68.0	97	14	H07566	H07566 crn124 BNL3
13	13.6	68.0	98	17	AZ608807	AZ608807 1M0433F13
c 14	13.4	67.0	90	9	AA066890	AA066890 mm10h12.r
15	13.2	66.0	27	14	R85601	R85601 yq28b07.s1
16	13.2	66.0	43	17	AZ776656	AZ776656 2M0010B04
17	13.2	66.0	80	9	AA822008	AA822008 vp25a12.r
c 18	13.2	66.0	88	17	TA25H04Q	AL453714 T. brucei
19	13.2	66.0	100	10	BE250488	BE250488 600943246
c 20	13.2	66.0	100	17	BH227999	BH227999 1006143E0
21	12.8	64.0	44	17	CNS07GIB	AL609797 Anopheles
c 22	12.8	64.0	71	17	AZ849964	AZ849964 2M0151E07
c 23	12.8	64.0	80	10	BE515127	BE515127 601236036
24	12.8	64.0	91	14	T72993	T72993 yc69a07.s1
25	12.8	64.0	91	17	HSJ227825	AJ227825 Homo sapi
c 26	12.8	64.0	96	13	BI711332	BI711332 id94b12.x
c 27	12.8	64.0	100	14	BQ164346	BQ164346 1091019B1
c 28	12.6	63.0	40	17	AZ592228	AZ592228 1M0403D02
29	12.6	63.0	49	13	BI442689	BI442689 dag58f11.
c 30	12.6	63.0	50	9	AU107274	AU107274 AU107274
c 31	12.6	63.0	50	9	AU107284	AU107284 AU107284
c 32	12.6	63.0	50	9	AU107288	AU107288 AU107288
c 33	12.6	63.0	54	17	AL753931	AL753931 Arabidops
c 34	12.6	63.0	74	10	AW638773	AW638773 bl74c06.w
c 35	12.6	63.0	75	17	AZ363321	AZ363321 1M0108J18
c 36	12.6	63.0	89	9	AA467218	AA467218 vd98h12.r
37	12.6	63.0	93	9	AA721402	AA721402 nz74a03.s
38	12.6	63.0	94	9	AA980283	AA980283 ua32f01.r
39	12.6	63.0	95	9	AI204853	AI204853 ZF-EST175
40	12.6	63.0	96	17	AZ764175	AZ764175 1M0560B06
41	12.6	63.0	98	9	AA911155	AA911155 oe69e02.s
c 42	12.6	63.0	100	10	AW313381	AW313381 618078B09
c 43	12.6	63.0	100	10	AW796888	AW796888 CM3-UM003
44	12.4	62.0	59	17	AZ507158	AZ507158 1M0348I07
c 45	12.4	62.0	67	17	AZ771369	AZ771369 1M0573B01

ALIGNMENTS

RESULT 1
AW454338
LOCUS AW454338 100 bp mRNA linear EST 17-FEB-2000
DEFINITION zeh11308 Zebrafish Embryonic Heart cDNA Library Danio rerio cDNA 5', mRNA sequence.
ACCESSION AW454338
VERSION AW454338.1 GI:6995125
KEYWORDS EST.
SOURCE zebrafish.
ORGANISM Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes ; Cyprinidae; Danio.
REFERENCE 1 (bases 1 to 100)
AUTHORS Ton,C., Mably,J.D., Dempsey,A.A., Hwang,D.M., Fishman,M.C. and Liew ,C.C.
TITLE Identification and Characterization of Expressed Sequence Tags from an Embryonic Zebrafish Heart cDNA Library
JOURNAL Unpublished (1999)
COMMENT Contact: Liew CC
Brigham and Women's Hospital
Harvard Medical School
75 Francis St. Boston, MA 02115, USA
Tel: 6177328915
Fax: 6179750995
Email: cliew@rics.bwh.harvard.edu
PCR PRimers
FORWARD: 5' GCCAAGCTCGAAATTAACCCTCACTAAAGGG 3'
BACKWARD: 5' CCAGTGAATTGTAATACGACTCACTATAGGGCG 3'

```
          Seq primer: 5' GAAATTAACCCTCACTAAAGGG 3'.
FEATURES             Location/Qualifiers
     source          1. .100
                     /organism="Danio rerio"
                     /db_xref="taxon:7955"
                     /clone_lib="Zebrafish Embryonic Heart cDNA Library"
                     /dev_stage="embryonic day 3 post-fertilization"
                     /lab_host="E.coli XL1-Blue mrF'"
                     /note="Organ: heart; Vector: Lambda ZAP Express; Site_1:
                     EcoRI; Site_2: XhoI; mRNA was purified from embryonic
                     zebrafish hearts (3 day post-fertilization). cDNA was
                     synthesized using a XhoI-Oligo dT adaptor-primer. EcoRI
                     adaptors were ligated, followed by digestion with XhoI,
                     for directional cloning into pre-digested lambda ZAP
                     Express vector. "
BASE COUNT       31 a     20 c     29 g     20 t
ORIGIN

  Query Match             74.0%;  Score 14.8;  DB 10;  Length 100;
  Best Local Similarity   88.9%;  Pred. No. 1.1e+04;
  Matches   16;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       2 CAACAGGCACAGTACAGC 19
           |||||  ||||||||||||
Db       3 CAACACTCACAGTACAGC 20


RESULT 2
AL632098/c
LOCUS       AL632098                 69 bp    mRNA    linear   EST 02-NOV-2001
DEFINITION  AL632098 XGC-gastrula Silurana tropicalis cDNA clone TGas016j18 5',
            mRNA sequence.
ACCESSION   AL632098
VERSION     AL632098.1  GI:16601581
KEYWORDS    EST.
SOURCE      western clawed frog.
  ORGANISM  Silurana tropicalis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Silurana.
REFERENCE   1  (bases 1 to 69)
  AUTHORS   Huckle,E., Taylor,R., Ashurst,J.L., Zorn,A.M. and Rogers,J.
  TITLE     Sanger Xenopus tropicalis EST project 2001 (10_2001)
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Huckle E
            Sanger Centre
            Hinxton, Cambridgeshire, CB10 1SA, UK
            Email: trop@sanger.ac.uk
            Sanger Xenopus tropicalis EST project 2001
            TROPICALIS_SEQUENCE_ID: TGas016j18.sp6
            Sequencing primer: SP6
            This sequence is from a Xenopus Gene Collection (XGC) library
            constructed by Aaron M. Zorn.
FEATURES             Location/Qualifiers
     source          1. .69
                     /organism="Silurana tropicalis"
                     /db_xref="taxon:8364"
                     /clone="TGas016j18"
                     /clone_lib="XGC-gastrula"
                     /dev_stage="gastrula (stages 10.5-13 mixed)"
                     /lab_host="Escherichia coli XL1-blue"
                     /note="Vector: pCS107; Site_1: EcoRI; Site_2: NotI; cDNA
                     was oligo dT primed from 5ug of poly A+ RNA from stages
                     10-13 gastrulae. EcoRI-NotI cut cDNA was then ligated
                     into pCS107 with EcoRI at the 5' end and NotI at the 3'
                     end."
BASE COUNT       10 a     16 c     14 g     29 t
ORIGIN

  Query Match             71.0%;  Score 14.2;  DB 9;  Length 69;
  Best Local Similarity   84.2%;  Pred. No. 1.8e+04;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 CAACAGGCACAGTACAGCA 20
           ||| || |||||||| |||
Db      55 CAATAGTCACAGTACTGCA 37


RESULT 3
AZ662265/c
LOCUS       AZ662265                 73 bp    DNA     linear   GSS 14-DEC-2000
DEFINITION  1M0541F15F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0541F15 F, DNA sequence.
ACCESSION   AZ662265
VERSION     AZ662265.1  GI:11799411
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 73)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0541  row: F  column: 15
            Seq primer: CGTTGTAAAACGACGGCCAGT
            Class: plasmid ends
            High quality sequence stop: 73.
FEATURES             Location/Qualifiers
     source          1. .73
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC1M0541F15"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
BASE COUNT       17 a     19 c     17 g     20 t
ORIGIN

  Query Match             71.0%;  Score 14.2;  DB 17;  Length 73;
  Best Local Similarity   84.2%;  Pred. No. 1.8e+04;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;
```

```
Qy      2 CAACAGGCACAGTACAGCA 20
          | |||||||||| | |||
Db     31 CCACAGGCACAGTTCTGCA 13


RESULT 4
AW477033/c
LOCUS       AW477033                 78 bp    mRNA    linear   EST 24-FEB-2000
DEFINITION  ga40h10.y1 Moss EST library PPU Physcomitrella patens cDNA clone
            PEP_SOURCE_ID:PPU041620 5', mRNA sequence.
ACCESSION   AW477033
VERSION     AW477033.1  GI:7047139
KEYWORDS    EST.
SOURCE      Physcomitrella patens.
  ORGANISM  Physcomitrella patens
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Bryophyta;
            Bryopsida; Funariidae; Funariales; Funariaceae; Physcomitrella.
REFERENCE   1  (bases 1 to 78)
  AUTHORS   Quatrano,R., Bashiardes,S., Cove,D., Cuming,A., Knight,C., Clifton
            ,S., Marra,M., Hillier,L., Pape,D., Martin,J., Wylie,T., Underwood
            ,K., Theising,B., Allen,M., Bowers,Y., Person,B., Swaller,T.,
            Steptoe,M., Gibbons,M., Harvey,N., Ritter,E., Jackson,Y., McCann,R.
            , Waterston,R. and Wilson,R.
  TITLE     Leeds/Wash U Moss EST Project
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Ralph Quatrano
            Leeds/Wash U Moss EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Libraries were constructed by Dr. Stavros Bashiardes as part of the
            Physcomitrella EST program (PEP) at the Univ. of Leeds (UK) and
            Washington Univ. in St. Louis (USA) DNA sequencing by: Washington
            University Genome Sequencing Center For information on obtaining a
            clone please contact: Celia Knight (c.d.knight@leeds.ac.uk)
            Seq primer: -40RP from Gibco
            High quality sequence stop: 71.
FEATURES             Location/Qualifiers
     source          1. .78
                     /organism="Physcomitrella patens"
                     /db_xref="taxon:3218"
                     /clone="PEP_SOURCE_ID:PPU041620"
                     /clone_lib="Moss EST library PPU"
                     /tissue_type="protonemata: 7 day old tissue
                     ammonium-grown"
                     /lab_host="DH10B"
                     /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
                     XhoI; Construction of the cDNA library was carried out
                     using Stratagenes 'UniZAP - cDNA synthesis kit'. cDNA
                     was constructed using an oligo dT primer/linker that
                     contains a XhoI site within it. Following ds cDNA
                     synthesis, EcoRI adapters were ligated to the blunt ends
                     and sample was digested with XhoI. The result is cDNA
                     with an EcoRI sticky end on one side and a XhoI sticky
                     end on the other. This cDNA was ligated directionally in
                     UniZAP arms. The vector is designed containing the
                     pBluescript sequence as well as lambda DNA and cDNA is
                     cloned within this pBluescript sequence. The vector was
                     then packaged using Gold gigapackaging extracts. Library
                     was grown in XLIBlue MRF' cells and amplified. The library
                     was excised by mass excision using Stratagens 'Mass
                     excision kit' that uses exassist as a helper phage that
                     releases the pBluescript sequence and circularises it as
                     single stranded plasmids that are then packaged (by helper
                     phage) and secreted out of the host cell as phagemids.
                     SOLR cells were transformed with phagemids and the library
                     was plated out on LB-amp plates to select for
                     transformants. Approximately 1,000,000 colonies were grown
                     and recovered. The double stranded plasmid library was
                     recovered by using Quiagen Midi prep kit. 2 micro grams of
                     each library were used to transform DH10B cells by
                     electroporation."
BASE COUNT       18 a     18 c     21 g     21 t
ORIGIN

  Query Match             71.0%;  Score 14.2;  DB 10;  Length 78;
  Best Local Similarity   84.2%;  Pred. No. 1.9e+04;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 CAACAGGCACAGTACAGCA 20
          || || ||||||| |||||
Db     20 CAGCATGCACAGTCCAGCA 2


RESULT 5
AV914039/c
LOCUS       AV914039                 71 bp    mRNA    linear   EST 18-JAN-2002
DEFINITION  AV914039 K. Sato unpublished cDNA library, cv. Haruna Nijo
            germination shoots Hordeum vulgare subsp. vulgare cDNA clone
            bags4e21 5', mRNA sequence.
ACCESSION   AV914039
VERSION     AV914039.1  GI:18209816
KEYWORDS    EST.
SOURCE      Hordeum vulgare subsp. vulgare.
  ORGANISM  Hordeum vulgare subsp. vulgare
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
            ; Triticeae; Hordeum.
REFERENCE   1  (bases 1 to 71)
  AUTHORS   Sato,K., Saisho,D. and Takeda,K.
  TITLE     Barley EST sequencing project in NIG and Okayama Univ
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Tadasu Shin-i
            Center For Genetic Resource Information
            National Institute of Genetics
            1111 Yata, Mishima, Shizuoka 411-8540, Japan
            Tel: 81-559-81-6856
            Fax: 81-559-81-6855
            Email: tshini@genes.nig.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .71
                     /organism="Hordeum vulgare subsp. vulgare"
                     /cultivar="Haruna Nijo"
                     /db_xref="taxon:112509"
                     /clone="bags4e21"
                     /clone_lib="K. Sato unpublished cDNA library, cv. Haruna
                     Nijo germination shoots"
                     /tissue_type="shoots"
                     /dev_stage="germination"
BASE COUNT        9 a     16 c     23 g     22 t      1 others
ORIGIN

  Query Match             70.0%;  Score 14;  DB 10;  Length 71;
  Best Local Similarity   100.0%;  Pred. No. 2.2e+04;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 CCAACAGGCACAGT 14
          ||||||||||||||
Db     14 CCAACAGGCACAGT 1


RESULT 6
AZ779007/c
LOCUS       AZ779007                 51 bp    DNA     linear   GSS 16-FEB-2001
DEFINITION  2M0014H16R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC2M0014H16 R, DNA sequence.
ACCESSION   AZ779007
VERSION     AZ779007.1  GI:12909227
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 51)
```

```
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000) .
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0014  row: H  column: 16
            Seq primer: CACACAGGAAACAGCTATGACC
            Class: plasmid ends
            High quality sequence stop: 51.
FEATURES              Location/Qualifiers
     source           1. .51
                      /organism="Mus musculus"
                      /strain="C57BL/6J"
                      /db_xref="taxon:10090"
                      /clone="UUGC2M0014H16"
                      /clone_lib="Mouse 10kb plasmid UUGC1M library"
                      /sex="Male"
                      /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                      /note="Vector: PWD42nv; Purified genomic DNA from M.
                      musculus C57BL/6J (male) was obtained  from the Jackson
                      Laboratory Mouse DNA Resource
                      (http://www.jax.org/resources/documents/dnares/). The DNA
                      was hydrodynamically sheared by repeated passage through a
                      0.005 inch orifice at constant velocity. The sheared DNA
                      was blunt end-repaired with T4 DNA polymerase and T4
                      polynucleotide kinase. Adaptor oligonucleotides were
                      ligated to the blunt ends in high  molar excess. The
                      adaptored DNA was purified and size-selected for a 9.5 to
                      10.5 kb range using preparative agarose gel
                      electrophoresis. Vector DNA was prepared from a derivative
                      of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                      inducible derivative of plasmid R1. The vector was ligated
                      with adaptors complementary to the insert adaptors and
                      purified. The sheared, adaptored mouse DNA was annealed to
                      adaptored vector DNA, and transformed into
                      chemically-competent E. coli XL10-Gold (Stratagene) cells
                      and selected for ampicillin resistance."
BASE COUNT       13 a     13 c     13 g     12 t
ORIGIN

  Query Match             69.0%;  Score 13.8;  DB 17;  Length 51;
  Best Local Similarity   88.2%;  Pred. No. 2.5e+04;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      2 CAACAGGCACAGTACAG 18
          |||||||||| | |||
Db     47 CAACAGGCACACTGCAG 31


RESULT 7
BQ760503/c
LOCUS       BQ760503                 86 bp    mRNA    linear   EST 26-JUL-2002
DEFINITION  EBro02_SQ005_N05_R root, 3 week, hydroponic grown, low nitrogen, cv
            Optic, EBro02 Hordeum vulgare cDNA clone EBro02_SQ005_N05 5', mRNA
            sequence.
ACCESSION   BQ760503
VERSION     BQ760503.1  GI:21968975
KEYWORDS    EST.
SOURCE      Hordeum vulgare.
  ORGANISM  Hordeum vulgare
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
            ; Triticeae; Hordeum.
REFERENCE   1  (bases 1 to 86)
  AUTHORS   Hedley,P., Liu,H., Caldwell,D., McCallum,N., Mudie,S., Cardle,L.,
            Ramsay,L., Machray,G., Marshall,D.F.M. and Waugh,R.
  TITLE     Development of Barley Transcriptome Resources
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Waugh R, Marshall DF
            Genome Dynamics/Computational Biology
            Scottish Crop Research Institute
            Invergowrie, Dundee, DD2 5DA, Scotland, UK
            Tel: 00 44 1382 562731
            Fax: 00 44 1382 562426
            Email: est@scri.sari.ac.uk.
FEATURES              Location/Qualifiers
     source           1. .86
                      /organism="Hordeum vulgare"
                      /cultivar="Optic"
                      /db_xref="taxon:4513"
                      /clone="EBro02_SQ005_N05"
                      /clone_lib="root, 3 week, hydroponic grown, low nitrogen,
                      cv Optic, EBro02"
                      /tissue_type="root"
                      /dev_stage="3 week"
                      /lab_host="DH10B"
                      /note="Vector: pSPORT1; Site_1: Sal I; Site_2: Not I;
                      Non-normalised library, directionally cloned into pSPORT1.
                      Derived from roots of 3 week old Nitrogen stressed barley
                      plants. Developed as part of the barley transcriptome
                      resources of BBSRC/SEERAD funded cereal IGF (Investigating
                      Gene Function) project."
BASE COUNT       16 a     19 c     16 g     35 t
ORIGIN

  Query Match             69.0%;  Score 13.8;  DB 14;  Length 86;
  Best Local Similarity   88.2%;  Pred. No. 2.9e+04;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      3 AACAGGCACAGTACAGC 19
          |||| ||| ||||||||
Db     27 AACATGCAAAGTACAGC 11


RESULT 8
D18618
LOCUS       D18618                   91 bp    mRNA    linear   EST 12-DEC-1995
DEFINITION  MUSGS01679 Mouse 3'-directed Mus musculus domesticus cDNA clone
            md1631 3', mRNA sequence.
ACCESSION   D18618
VERSION     D18618.1  GI:1100587
KEYWORDS    EST.
SOURCE      western European house mouse.
  ORGANISM  Mus musculus domesticus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 91)
  AUTHORS   Kawamoto,S., Okubo,K., Yoshii,J., Katsuki,M. and Matsubara,K.
  TITLE     Analysis of gene expression in mouse embryogenesis by 3'-directed
            cDNA sequencing
  JOURNAL   Unpublished (1995)
COMMENT     Contact: Kawamoto,S., Okubo,K., Yoshii,J., Katsuki,M. and Matsubara
            ,K.
            Institute for Cellular and Molecular Biology
            Osaka University
            3-1 Yamada-oka, Suita, Osaka 565, Japan.
FEATURES              Location/Qualifiers
     source           1. .91
                      /organism="Mus musculus domesticus"
                      /strain="C57BL/6J"
                      /db_xref="taxon:10092"
                      /clone="md1631"
                      /clone_lib="Mouse 3'-directed"
                      /tissue_type="decidual tissue (day 6.5-8.5 of gestation)"
BASE COUNT       27 a     19 c     23 g     19 t      3 others
```

```
ORIGIN

  Query Match            69.0%;  Score 13.8;  DB 14;  Length 91;
  Best Local Similarity  88.2%;  Pred. No. 2.9e+04;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       4 ACAGGCACAGTACAGCA 20
           |||| ||||| ||||||
Db      35 ACAGCCACAGGACAGCA 51


RESULT 9
CNS02A9J
LOCUS       CNS02A9J                  94 bp    DNA     linear   GSS 12-MAY-2000
DEFINITION  Tetraodon nigroviridis genome survey sequence PUC-Ori end of clone
            251E02 of library G from Tetraodon nigroviridis, genomic survey
            sequence.
ACCESSION   AL188272
VERSION     AL188272.1  GI:7826376
KEYWORDS    GSS; genome survey sequence.
SOURCE      Tetraodon nigroviridis.
  ORGANISM  Tetraodon nigroviridis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
            Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
            Tetraodontidae; Tetraodon.
REFERENCE   1  (bases 1 to 94)
  AUTHORS   Roest-Crollius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C.,
            Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F.,
            Saurin,W. and Weissenbach,J.
  TITLE     Human gene number estimate provided by genome wide analysis using
            Tetraodon nigroviridis DNA sequence
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 94)
  AUTHORS   Roest-Crollius,H., Jaillon,O., Dasilva,C., Fizames,C., Fisher,C.,
            Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and
            Weissenbach,J.
  TITLE     Charaterization and repeat analysis of the compact genome of the
            freshwater pufferfish Tetraodon nigroviridis
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 94)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-APR-2000)
COMMENT     This sequence is a single read and was generated as part of a large
            scale clone-end sequencing project of the Tetraodon nigroviridis
            genome. For more information, please take a look at
            http://www.genoscope.cns.fr/Tetraodon.
FEATURES             Location/Qualifiers
     source          1. .94
                     /organism="Tetraodon nigroviridis"
                     /db_xref="taxon:99883"
                     /clone="251E02"
                     /clone_lib="G"
                     /note="Genoscope sequence ID : C0AG251BC01SP1~end :
                     PUC-Ori"
BASE COUNT       29 a     24 c     27 g     12 t      2 others
ORIGIN

  Query Match            69.0%;  Score 13.8;  DB 17;  Length 94;
  Best Local Similarity  88.2%;  Pred. No. 3e+04;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 CCAACAGGCACAGTACA 17
           ||||||||||| || |
Db      63 CCAACAGGCACACTAAA 79


RESULT 10
BQ807426/c
LOCUS       BQ807426                  87 bp    mRNA    linear   EST 31-JUL-2002
DEFINITION  NISC_kk03g08.y2 NCI_CGAP_Brn72 Macaca mulatta cDNA clone
            IMAGE:5330510 5', mRNA sequence.
ACCESSION   BQ807426
VERSION     BQ807426.1  GI:22031635
KEYWORDS    EST.
SOURCE      rhesus monkey.
  ORGANISM  Macaca mulatta
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
            Cercopithecinae; Macaca.
REFERENCE   1  (bases 1 to 87)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            cDNA Library Preparation:
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium/LLNL
             DNA Sequencing by: National Institutes of Health Intramural
            Sequencing Center (NISC)
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            info@image.llnl.gov
            Plate: LLAM11838  row: N  column: 15
            Seq primer: M13RP1 reverse primer (ABI).
FEATURES             Location/Qualifiers
     source          1. .87
                     /organism="Macaca mulatta"
                     /db_xref="taxon:9544"
                     /clone="IMAGE:5330510"
                     /clone_lib="NCI_CGAP_Brn72"
                     /tissue_type="hypothalamus"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: brain; Vector: pCMV-SPORT6.ccdb; Site_1:
                     NotI; Site_2: EcoRV; Cloned unidirectionally. Primer:
                     Oligo dT. Average insert size 2.2 kb. Constructed by
                     Invitrogen. Note: this is a NCI_CGAP Library."
BASE COUNT       14 a     23 c     20 g     30 t
ORIGIN

  Query Match            68.0%;  Score 13.6;  DB 14;  Length 87;
  Best Local Similarity  80.0%;  Pred. No. 3.5e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 CCAACAGGCACAGTACAGCA 20
           | ||||||| ||| ||| ||
Db      71 CAAACAGGCCCAGCACATCA 52


RESULT 11
F32010
LOCUS       F32010                    90 bp    mRNA    linear   EST 13-MAY-1999
DEFINITION  HSPD23880 HM3 Homo sapiens cDNA clone s4000106E10, mRNA sequence.
ACCESSION   F32010
VERSION     F32010.1  GI:4817636
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 90)
  AUTHORS   Lanfranchi,G., Muraro,T., Caldara,F., Pacchioni,B., Pallavicini,A.,
            Pandolfo,D., Toppo,S., Trevisan,S., Scarso,S. and Valle,G.
  TITLE     Identification of 4370 expressed sequence tags from a
            3'-end-specific cDNA library of human skeletal muscle by DNA
            sequencing and filter hybridization
  JOURNAL   Genome Res. 6 (1), 35-42 (1996)
  MEDLINE   96276048
COMMENT     Contact: Valle G.
            CRIBI Biotechnology Centre
            University of Padua
            Via Trieste 75, 35121 Padua, Italy
            ABI Chromatograms and other information are available on WWW at
            http://grup.bio.unipd.it.
```

```
FEATURES             Location/Qualifiers
     source          1. .90
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="s4000106E10"
                     /clone_lib="HM3"
                     /sex="female"
                     /tissue_type="pectoral muscle (after mastectomy)"
                     /note="Vector: pcDNAII (Invitrogen); Site_1: BstXI;
                     Site_2: NotI; The library was constructed by G.
                     Lanfranchi. This library is not subtracted nor normalized.
                     The first strand cDNA was primed with a biotinylated
                     oligo-dT-NotI primer
                     (5'-biotin-AACCCGGCTCGAGCGGCCGCTTTTTTTTTTTTTTTT-3'). The
                     ds cDNA was sonicated and size-selected in the range
                     350-550  bp. The 3' specific fragments were selected by
                     streptavidin coated magnetic beads, ligated to
                     non-palindromic BstXI adapters, NotI digested and
                     directionally cloned into BstXI-NotI cut pcDNAII vector."
BASE COUNT       21 a     20 c     27 g     22 t
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 14;  Length 90;
  Best Local Similarity   80.0%;  Pred. No. 3.6e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 CCAACAGGCACAGTACAGCA 20
           ||| | | ||||||| ||||
Db      68 CCACCTGCCACAGTAAAGCA 87


RESULT 12
H07566/c
LOCUS       H07566                   97 bp    mRNA    linear   EST 23-JUN-1995
DEFINITION  crn124 BNL3 Brassica napus cDNA 5', mRNA sequence.
ACCESSION   H07566
VERSION     H07566.1  GI:872390
KEYWORDS    EST.
SOURCE      rape.
  ORGANISM  Brassica napus
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE   1  (bases 1 to 97)
  AUTHORS   Sohn,U., Lee,C.M., Cho,K.H., Kim,N.I., Cho,T.J., Kwak,J.M. and Nam
            ,H.G.
  TITLE     Brassica napus (rape) cDNAs
  JOURNAL   Unpublished (1995)
COMMENT     Contact: Uik Sohn
            Laboratory of Molecular Biology
            Kyungpook National University
            Dept. of Genetic Eng.,Kyungpook National Univ.,Taegu 702-701,Korea
            Tel: 0539505382
            Fax: 0539555327
            Email: usohn@bh.kyungpook.ac.kr
            EST is putatively homologous to unknown gene.
            Seq primer: M13 forward/reverse.
FEATURES             Location/Qualifiers
     source          1. .97
                     /organism="Brassica napus"
                     /strain="cv. Naehan"
                     /db_xref="taxon:3708"
                     /clone_lib="BNL3"
                     /lab_host="JM109"
                     /note="Vector: M13mp19; Site_1: HincII; Site_2: EcoRI;
                     cDNA library was constructed by oligo(dT) priming and
                     directionally  cloned into HincII/EcoRI site of M13mp19. "
BASE COUNT       24 a     19 c     34 g     20 t
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 14;  Length 97;
  Best Local Similarity   80.0%;  Pred. No. 3.7e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 CCAACAGGCACAGTACAGCA 20
           | |||||||| ||| || ||
Db      24 CGAACAGGCATAGTTCACCA 5


RESULT 13
AZ608807
LOCUS       AZ608807                 98 bp    DNA     linear   GSS 13-DEC-2000
DEFINITION  1M0433F13F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0433F13 F, DNA sequence.
ACCESSION   AZ608807
VERSION     AZ608807.1  GI:11730997
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 98)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0433  row: F  column: 13
            Seq primer: CGTTGTAAAACGACGGCCAGT
            Class: plasmid ends
            High quality sequence stop: 98.
FEATURES             Location/Qualifiers
     source          1. .98
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC1M0433F13"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
BASE COUNT       35 a     20 c     19 g     24 t
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 17;  Length 98;
  Best Local Similarity   80.0%;  Pred. No. 3.7e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;
```

```
Qy      1 CCAACAGGCACAGTACAGCA 20
          | |||||||| ||| || ||
Db     19 CTAACAGGCAGAGTCCATCA 38


RESULT 14
AA066890/c
LOCUS       AA066890                 90 bp    mRNA    linear   EST 04-FEB-1997
DEFINITION  mm10h12.r1 Stratagene mouse diaphragm (#937303) Mus musculus cDNA
            clone IMAGE:521159 5' similar to gb:Y00371_rna1 HEAT SHOCK COGNATE
            71 KD PROTEIN (HUMAN); gb:U27129 Mus musculus breast heat shock 73
            protein (MOUSE);, mRNA sequence.
ACCESSION   AA066890
VERSION     AA066890.1  GI:1564689
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 90)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:315007
            Seq primer: -28m13 rev1 ET from Amersham
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .90
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:521159"
                     /clone_lib="Stratagene mouse diaphragm (#937303)"
                     /tissue_type="diaphragm"
                     /dev_stage="adult"
                     /lab_host="SOLR (kanamycin resistant)"
                     /note="Organ: diaphragm; Vector: pBluescript SK-; Site_1:
                     EcoRI; Site_2: XhoI; Cloned unidirectionally from mRNA
                     prepared from diaphragm muscle.  Primer: Oligo dT. Average
                     insert size: 1.5 kb. Uni-ZAP XR Vector; -5' adaptor
                     sequence: 5' GAATTCGGCACGAG 3' -3' adaptor sequence: 5'
                     CTCGAGTTTTTTTTTTTTTTTTT 3'"
BASE COUNT       19 a     17 c     25 g     29 t
ORIGIN

  Query Match             67.0%;  Score 13.4;  DB 9;  Length 90;
  Best Local Similarity   93.3%;  Pred. No. 4.4e+04;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      6 AGGCACAGTACAGCA 20
          |||||||||||| ||
Db     56 AGGCACAGTACATCA 42


RESULT 15
R85601
LOCUS       R85601                   27 bp    mRNA    linear   EST 22-JAN-1999
DEFINITION  yq28b07.s1 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone
            IMAGE:274908 3' similar to gb:J05249 REPLICATION PROTEIN A 32 KD
            SUBUNIT (HUMAN);, mRNA sequence.
ACCESSION   R85601
VERSION     R85601.1  GI:944007
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 27)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman
            ,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J.,
            Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston
            ,R., Williamson,A., Wohldmann,P. and Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Insert Size: 713
            High quality sequence starts: 1
            High quality sequence stops: 1
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Trace considered overall poor quality
            Insert Length: 713   Std Error: 0.00
            Seq primer: Promega -21m13
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .27
                     /organism="Homo sapiens"
                     /db_xref="GDB:3798931"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:274908"
                     /clone_lib="Soares fetal liver spleen 1NFLS"
                     /sex="male"
                     /dev_stage="20 week-post conception fetus"
                     /lab_host="DH10B (ampicillin resistant)"
                     /note="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia)
                     with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
                     1st strand cDNA was primed with a Pac I - oligo(dT) primer
                     [5' AACTGGAAGAATTAATTAAAGATCTTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Pac I and cloned into the Pac I
                     and Eco RI sites of the modified pT7T3 vector.  Library
                     went through one round of normalization. Library
                     constructed by Bento Soares and M.Fatima Bonaldo."
BASE COUNT       12 a      5 c      9 g      1 t
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 14;  Length 27;
  Best Local Similarity   83.3%;  Pred. No. 3.8e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      3 AACAGGCACAGTACAGCA 20
          |||||||| ||  |||||
Db      8 AACAGGCAGAGGCCAGCA 25


Search completed: November 26, 2002, 17:58:22
Job time : 809.5 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 3, 2002, 18:14:22 ; Search time 351.3 Seconds
(without alignments)
1656.863 Million cell updates/sec

Title: US-09-296-264-29
Perfect score: 20
Sequence: 1 accatccacaagttcaaagt 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 995600

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		14.8	74.0	30	6	A46141	A46141 Sequence 36
2		14.8	74.0	30	6	AR196867	AR196867 Sequence
3		14.2	71.0	24	6	AR200099	AR200099 Sequence
4		14.2	71.0	28	6	AR161706	AR161706 Sequence
5	c	13.8	69.0	22	6	AX050099	AX050099 Sequence
6		13.8	69.0	30	6	AR122567	AR122567 Sequence
7		13.4	67.0	43	6	AX112273	AX112273 Sequence
8		13.2	66.0	51	6	AX304319	AX304319 Sequence
9	c	13.2	66.0	51	6	AX304320	AX304320 Sequence
10		13.2	66.0	65	6	AX486002	AX486002 Sequence
11	c	12.8	64.0	22	6	AX486892	AX486892 Sequence
12	c	12.8	64.0	30	6	AX455520	AX455520 Sequence
13	c	12.8	64.0	33	6	AR010280	AR010280 Sequence
14	c	12.8	64.0	33	6	AR026173	AR026173 Sequence
15		12.8	64.0	33	6	AR026253	AR026253 Sequence
16	c	12.8	64.0	33	6	AR026290	AR026290 Sequence
17		12.8	64.0	33	6	AR026294	AR026294 Sequence
18	c	12.8	64.0	33	6	AR050279	AR050279 Sequence
19	c	12.8	64.0	33	6	I61377	I61377 Sequence 6
20	c	12.8	64.0	33	6	I81312	I81312 Sequence 13
21	c	12.8	64.0	33	6	I82915	I82915 Sequence 17
22		12.8	64.0	33	6	I82995	I82995 Sequence 97
23	c	12.8	64.0	33	6	I83032	I83032 Sequence 13
24		12.8	64.0	33	6	I83036	I83036 Sequence 13
25		12.8	64.0	41	6	AR109075	AR109075 Sequence
26		12.8	64.0	41	6	AR200730	AR200730 Sequence
27	c	12.8	64.0	42	6	AX322473	AX322473 Sequence
28	c	12.8	64.0	47	6	AX322469	AX322469 Sequence
29	c	12.8	64.0	51	6	AX165608	AX165608 Sequence
30	c	12.8	64.0	52	6	AX322465	AX322465 Sequence
31		12.8	64.0	61	9	HUMFAH09	L14666 Human fumar
32	c	12.6	63.0	19	6	AX148017	AX148017 Sequence
33		12.6	63.0	27	9	HUMKRT14	D28807 Homo sapien
34	c	12.6	63.0	28	6	AR090955	AR090955 Sequence
35	c	12.6	63.0	28	6	AR197990	AR197990 Sequence
36		12.6	63.0	29	6	AR183285	AR183285 Sequence
37		12.6	63.0	29	6	AX006532	AX006532 Sequence
38		12.6	63.0	29	6	AX006636	AX006636 Sequence
39		12.6	63.0	29	6	AX009643	AX009643 Sequence
40		12.6	63.0	29	6	AX010905	AX010905 Sequence
41		12.6	63.0	29	6	AX030240	AX030240 Sequence
42		12.6	63.0	29	6	AX034826	AX034826 Sequence
43		12.6	63.0	29	6	AX049920	AX049920 Sequence
44		12.6	63.0	29	6	AX076413	AX076413 Sequence
45		12.6	63.0	29	6	AX077162	AX077162 Sequence

ALIGNMENTS

RESULT 1
A46141
LOCUS A46141 30 bp DNA linear PAT 07-MAR-1997
DEFINITION Sequence 36 from Patent WO9522626.
ACCESSION A46141
VERSION A46141.1 GI:2300389
KEYWORDS .
SOURCE unidentified.
ORGANISM unidentified
unclassified.
REFERENCE 1 (bases 1 to 30)
AUTHORS Meijer,C.J., Van,D.B., Walboomers,J.M. and Snijders,P.J.
TITLE HUMAN PAPILLOMA VIRUS DETECTION IN A NUCLEIC ACID AMPLIFICATION PROCESS USING GENERAL PRIMERS
JOURNAL Patent: WO 9522626-A 36 24-AUG-1995;

```
            STICHTING RESEARCHFONDS PATHOL (NL)
COMMENT     Other publication AU 1672295 950904.
FEATURES             Location/Qualifiers
     source          1. .30
                     /organism="unidentified"
                     /db_xref="taxon:32644"
BASE COUNT       14 a      9 c      2 g .    5 t
ORIGIN

  Query Match             74.0%;  Score 14.8;  DB 6;  Length 30;
  Best Local Similarity   88.9%;  Pred. No. 4.7e+03;
  Matches   16;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 ACCATCCACAAGTTCAAA 18
          || |||||||||| ||||
Db      4 ACAATCCACAAGTACAAA 21


RESULT 2
AR196867
LOCUS       AR196867                  30 bp    DNA     linear   PAT 20-APR-2002
DEFINITION  Sequence 36 from patent US 6352825.
ACCESSION   AR196867
VERSION     AR196867.1  GI:20246716
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 30)
  AUTHORS   Meijer,C.Joannes., van den Brule,A.Johannes., Walboomers,J.Marcus.
            and Snijders,P.Josephus.
  TITLE     Human Papilloma Virus detection in a nucleic acid amplification
            process using general primers
  JOURNAL   Patent: US 6352825-A 36 05-MAR-2002;
FEATURES             Location/Qualifiers
     source          1. .30
                     /organism="unknown"
BASE COUNT       14 a      9 c      2 g      5 t
ORIGIN

  Query Match             74.0%;  Score 14.8;  DB 6;  Length 30;
  Best Local Similarity   88.9%;  Pred. No. 4.7e+03;
  Matches   16;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 ACCATCCACAAGTTCAAA 18
          || |||||||||| ||||
Db      4 ACAATCCACAAGTACAAA 21


RESULT 3
AR200099
LOCUS       AR200099                  24 bp    DNA     linear   PAT 20-APR-2002
DEFINITION  Sequence 10 from patent US 6355777.
ACCESSION   AR200099
VERSION     AR200099.1  GI:20250173
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 24)
  AUTHORS   Walker,D.H. and McBride,J.W.
  TITLE     P43 antigen for the immunodiagnosis of canine ehrlichiosis and uses
            thereof
  JOURNAL   Patent: US 6355777-A 10 12-MAR-2002;
FEATURES             Location/Qualifiers
     source          1. .24
                     /organism="unknown"
BASE COUNT       10 a      7 c      1 g      6 t
ORIGIN

  Query Match             71.0%;  Score 14.2;  DB 6;  Length 24;
  Best Local Similarity   84.2%;  Pred. No. 9.2e+03;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 CCATCCACAAGTTCAAAGT 20
          ||||| |||||| |||| |
Db      2 CCATCTACAAGTCCAAAAT 20


RESULT 4
AR161706
LOCUS       AR161706                  28 bp    DNA     linear   PAT 17-OCT-2001
DEFINITION  Sequence 16 from patent US 6258529.
ACCESSION   AR161706
VERSION     AR161706.1  GI:16228590
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 28)
  AUTHORS   Berdoz,J. and Kraehenbuhl,J.-P.
  TITLE     PCR amplification of rearranged genomic variable regions of
            immunoglobulin genes
  JOURNAL   Patent: US 6258529-A 16 10-JUL-2001;
FEATURES             Location/Qualifiers
     source          1. .28
                     /organism="unknown"
BASE COUNT       12 a      5 c      6 g      5 t
ORIGIN

  Query Match             71.0%;  Score 14.2;  DB 6;  Length 28;
  Best Local Similarity   84.2%;  Pred. No. 9.4e+03;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 ACCATCCACAAGTTCAAAG 19
          | | || ||||||||||||
Db      2 ATCGTCGACAAGTTCAAAG 20


RESULT 5
AX050099/c
LOCUS       AX050099                  22 bp    DNA     linear   PAT 13-JAN-2001
DEFINITION  Sequence 112 from Patent WO0071710.
ACCESSION   AX050099
VERSION     AX050099.1  GI:12226499
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 22)
  AUTHORS   Denefle,P., Rosier-Montus,M.F., Arnould-Reguigne,I., Prades,C. and
            Clepet,C.
  TITLE     Expression products of genes involved in diseases related to
            cholesterol metabolism
  JOURNAL   Patent: WO 0071710-A 112 30-NOV-2000;
            Aventis Pharma S.A. (FR)
FEATURES             Location/Qualifiers
     source          1. .22
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT        3 a      3 c      8 g      8 t
ORIGIN

  Query Match             69.0%;  Score 13.8;  DB 6;  Length 22;
  Best Local Similarity   88.2%;  Pred. No. 1.5e+04;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 ACCATCCACAAGTTCAA 17
          || ||||||| ||||||
Db     21 ACAATCCACAGGTTCAA 5


RESULT 6
AR122567
LOCUS       AR122567                  30 bp    DNA     linear   PAT 16-MAY-2001
```

```
DEFINITION  Sequence 10 from patent US 6165737.
ACCESSION   AR122567
VERSION     AR122567.1  GI:14106884
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 30)
  AUTHORS   Wang,X. and Liu,X.
  TITLE     DNA fragmentation factor involved in apoptosis
  JOURNAL   Patent: US 6165737-A 10 26-DEC-2000;
FEATURES             Location/Qualifiers
     source          1. .30
                     /organism="unknown"
BASE COUNT       13 a     11 c      3 g      3 t
ORIGIN

  Query Match             69.0%;  Score 13.8;  DB 6;  Length 30;
  Best Local Similarity   88.2%;  Pred. No. 1.5e+04;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 ACCATCCACAAGTTCAA 17
          |||| ||||||| ||||
Db     13 ACCACCCACAAGCTCAA 29


RESULT 7
AX112273
LOCUS       AX112273                 43 bp    DNA     linear   PAT 01-MAY-2001
DEFINITION  Sequence 2 from Patent WO0127304.
ACCESSION   AX112273
VERSION     AX112273.1  GI:13939079
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 43)
  AUTHORS   Charneau,P., Zennou,V., Pflumio,F., Sirven,A. and Dubart
            Kupperschmitt,A.
  TITLE     Lentiviral triplex dna, and vectors and recombinant cells
            containing lentiviral triplex dna
  JOURNAL   Patent: WO 0127304-A 2 19-APR-2001;
            INSTITUT PASTEUR (FR) ; INSTITUT NATIONAL DE LA SANTE ET DE LA
            RECHERCHE MEDICALE (INSERM) (FR)
FEATURES             Location/Qualifiers
     source          1. .43
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="MUTAGENESIS PRIMER BASED ON PLASMID pLAI3"
BASE COUNT        9 a     11 c     13 g     10 t
ORIGIN

  Query Match             67.0%;  Score 13.4;  DB 6;  Length 43;
  Best Local Similarity   93.3%;  Pred. No. 2.6e+04;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      3 CATCCACAAGTTCAA 17
          ||||||||| |||||
Db      4 CATCCACAACTTCAA 18


RESULT 8
AX304319
LOCUS       AX304319                 51 bp    DNA     linear   PAT 30-NOV-2001
DEFINITION  Sequence 1127 from Patent WO0183525.
ACCESSION   AX304319
VERSION     AX304319.1  GI:17383776
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Feige,U., Liu,C.F., Cheetham,J.C., Boone,T.C. and Gudas,J.M.
  TITLE     Modified peptides as therapeutic agents
  JOURNAL   Patent: WO 0183525-A 1127 08-NOV-2001;
            Amgen Inc., (US)
FEATURES             Location/Qualifiers
     source          1. .51
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="ANTI-SENSE PCR PRIMER FOR VEGF ANTAGONIST
                     CONSTRUCT"
BASE COUNT       15 a     17 c      7 g     12 t
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 6;  Length 51;
  Best Local Similarity   83.3%;  Pred. No. 3.4e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 ACCATCCACAAGTTCAAA 18
          |||| ||| | |||||||
Db     28 ACCACCCAGACGTTCAAA 45


RESULT 9
AX304320/c
LOCUS       AX304320                 51 bp    DNA     linear   PAT 30-NOV-2001
DEFINITION  Sequence 1128 from Patent WO0183525.
ACCESSION   AX304320
VERSION     AX304320.1  GI:17383777
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Feige,U., Liu,C.F., Cheetham,J.C., Boone,T.C. and Gudas,J.M.
  TITLE     Modified peptides as therapeutic agents
  JOURNAL   Patent: WO 0183525-A 1128 08-NOV-2001;
            Amgen Inc., (US)
FEATURES             Location/Qualifiers
     source          1. .51
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="SENSE PCR PRIMER FOR Fc CONSTRUCT"
BASE COUNT       12 a      7 c     17 g     15 t
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 6;  Length 51;
  Best Local Similarity   83.3%;  Pred. No. 3.4e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 ACCATCCACAAGTTCAAA 18
          |||| ||| | |||||||
Db     24 ACCACCCAGACGTTCAAA 7


RESULT 10
AX486002
LOCUS       AX486002                 65 bp    DNA     linear   PAT 16-AUG-2002
DEFINITION  Sequence 3302 from Patent WO02053728.
ACCESSION   AX486002
VERSION     AX486002.1  GI:22320218
KEYWORDS    .
SOURCE      Candida albicans.
  ORGANISM  Candida albicans
            Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
            Saccharomycetales; mitosporic Saccharomycetales; Candida.
REFERENCE   1
  AUTHORS   Roemer,T., Jiang,B., Boone,C., Bussey,H. and Ohlsen,K.L.
  TITLE     Gene disruption methodologies for drug target discovery
  JOURNAL   Patent: WO 02053728-A 3302 11-JUL-2002;
            Elitra Pharmaceuticals, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .65
                     /organism="Candida albicans"
                     /db_xref="taxon:5476"
```

```
BASE COUNT       19 a     11 c     10 g     25 t
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 6;  Length 65;
  Best Local Similarity   83.3%;  Pred. No. 3.5e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 ACCATCCACAAGTTCAAA 18
          | ||||||||||||  ||
Db     24 ATCATCCACAAGTTTTAA 41


RESULT 11
AX486892/c
LOCUS       AX486892                  22 bp    DNA     linear   PAT 16-AUG-2002
DEFINITION  Sequence 4192 from Patent WO02053728.
ACCESSION   AX486892
VERSION     AX486892.1  GI:22321040
KEYWORDS    .
SOURCE      Candida albicans.
  ORGANISM  Candida albicans
            Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
            Saccharomycetales; mitosporic Saccharomycetales; Candida.
REFERENCE   1
  AUTHORS   Roemer,T., Jiang,B., Boone,C., Bussey,H. and Ohlsen,K.L.
  TITLE     Gene disruption methodologies for drug target discovery
  JOURNAL   Patent: WO 02053728-A 4192 11-JUL-2002;
            Elitra Pharmaceuticals, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .22
                     /organism="Candida albicans"
                     /db_xref="taxon:5476"
BASE COUNT        5 a      2 c      8 g      7 t
ORIGIN

  Query Match             64.0%;  Score 12.8;  DB 6;  Length 22;
  Best Local Similarity   87.5%;  Pred. No. 4.8e+04;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      5 TCCACAAGTTCAAAGT 20
          ||||||| ||||||  |
Db     22 TCCACAACTTCAAACT 7


RESULT 12
AX455520/c
LOCUS       AX455520                  30 bp    DNA     linear   PAT 06-JUL-2002
DEFINITION  Sequence 8 from Patent WO0218584.
ACCESSION   AX455520
VERSION     AX455520.1  GI:21714599
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Pastan,I.H., Liu,X.F., Bera,T.K., Lee,B. and Egland,K.A.
  TITLE     Xage-1, a gene expressed in multiple cancers, and uses thereof
  JOURNAL   Patent: WO 0218584-A 8 07-MAR-2002;
            THE SECRETARY OF THE DEPARTMENT OF HEALTH AND HUMAN SERVICES (US)
FEATURES             Location/Qualifiers
     source          1. .30
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="primer xa-1"
BASE COUNT        5 a      6 c      6 g     13 t
ORIGIN

  Query Match             64.0%;  Score 12.8;  DB 6;  Length 30;
  Best Local Similarity   87.5%;  Pred. No. 5.1e+04;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      3 CATCCACAAGTTCAAA 18
          || |||||||||  |||
Db     29 CAACCACAAGTTTAAA 14


RESULT 13
AR010280/c
LOCUS       AR010280                  33 bp    DNA     linear   PAT 04-DEC-1998
DEFINITION  Sequence 6 from patent US 5756709.
ACCESSION   AR010280
VERSION     AR010280.1  GI:3969085
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 33)
  AUTHORS   Nelson,N.Charles., Woodhead,J.Stuart., Weeks,I. and Cheikh,A.Ben.
  TITLE     Compositions for the simultaneous detection and quantification of
            multiple specific nucleic acid sequences
  JOURNAL   Patent: US 5756709-A 6 26-MAY-1998;
FEATURES             Location/Qualifiers
     source          1. .33
                     /organism="unknown"
BASE COUNT        7 a      9 c      4 g     13 t
ORIGIN

  Query Match             64.0%;  Score 12.8;  DB 6;  Length 33;
  Best Local Similarity   87.5%;  Pred. No. 5.1e+04;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      3 CATCCACAAGTTCAAA 18
          ||||||||| || |||
Db     33 CATCCACAATTTTAAA 18


RESULT 14
AR026173/c
LOCUS       AR026173                  33 bp    DNA     linear   PAT 29-SEP-1999
DEFINITION  Sequence 17 from patent US 5856088.
ACCESSION   AR026173
VERSION     AR026173.1  GI:5937013
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 33)
  AUTHORS   McDonough,S.H., Ryder,T.B. and Yang,Y.
  TITLE     Detection of human immunodeficiency virus type 1
  JOURNAL   Patent: US 5856088-A 17 05-JAN-1999;
FEATURES             Location/Qualifiers
     source          1. .33
                     /organism="unknown"
BASE COUNT        7 a      9 c      4 g     13 t
ORIGIN

  Query Match             64.0%;  Score 12.8;  DB 6;  Length 33;
  Best Local Similarity   87.5%;  Pred. No. 5.1e+04;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      3 CATCCACAAGTTCAAA 18
          ||||||||| || |||
Db     33 CATCCACAATTTTAAA 18


RESULT 15
AR026253
LOCUS       AR026253                  33 bp    DNA     linear   PAT 29-SEP-1999
DEFINITION  Sequence 97 from patent US 5856088.
ACCESSION   AR026253
VERSION     AR026253.1  GI:5937093
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 33)
```

```
  AUTHORS   McDonough,S.H., Ryder,T.B. and Yang,Y.
  TITLE     Detection of human immunodeficiency virus type 1
  JOURNAL   Patent: US 5856088-A 97 05-JAN-1999;
FEATURES             Location/Qualifiers
     source          1. .33
                     /organism="unknown"
BASE COUNT       13 a      4 c      9 g      7 t
ORIGIN

  Query Match             64.0%;  Score 12.8;  DB 6;  Length 33;
  Best Local Similarity   87.5%;  Pred. No. 5.1e+04;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       3 CATCCACAAGTTCAAA 18
           ||||||||| || |||
Db       1 CATCCACAATTTTAAA 16


Search completed: December  3, 2002, 22:23:51
Job time : 357.3 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 06:29:45 ; Search time 94.1 Seconds
(without alignments)
478.639 Million cell updates/sec

Title: US-09-296-264-29
Perfect score: 20
Sequence: 1 accatccacaagttcaaagt 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 2390332

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	20	100.0	20	21	AAZ31459	Human neuropilin m
	2	14.8	74.0	30	16	AAT00985	Human papilloma vi
	3	14.8	74.0	30	22	AAS14559	HPV oligonucleotid
c	4	14.8	74.0	84	22	AAC87390	Human prostate can
	5	14.4	72.0	65	24	ABN29990	Rat spliced transc
	6	14.2	71.0	28	17	AAT30466	Primer 5'kappa3 fo
	7	14.2	71.0	49	22	AAS11817	Mouse surfactant p
	8	14.2	71.0	91	24	ABS18070	Human genome-deriv
c	9	13.8	69.0	22	23	ABN93486	Human gene GS91301
	10	13.8	69.0	27	22	AAF25277	PCR primer used to
	11	13.8	69.0	30	21	AAZ38999	Human DNA fragment
	12	13.8	69.0	30	22	AAH74649	PCR primer for cDN
	13	13.8	69.0	30	22	AAD02505	Human DNA fragment
c	14	13.8	69.0	65	24	ABN28293	Rat spliced transc
	15	13.8	69.0	65	24	ABN55280	Mouse spliced tran
c	16	13.8	69.0	80	22	ABA51062	Human breast cell
c	17	13.8	69.0	80	22	ABA69052	Human foetal liver
c	18	13.8	69.0	80	22	ABA35995	Probe #14461 for g
c	19	13.8	69.0	80	22	AAK17364	Human brain expres
c	20	13.8	69.0	80	22	AAK43160	Human bone marrow
c	21	13.8	69.0	80	22	AAI23933	Probe #13866 for g
c	22	13.8	69.0	80	22	AAI49240	Probe #17926 used
c	23	13.8	69.0	80	22	AAI09533	Probe #9524 used t
c	24	13.8	69.0	80	24	ABS17237	Human genome-deriv
	25	13.8	69.0	99	17	AAT30893	Primer 23 for 95 k
c	26	13.6	68.0	60	24	ABN50173	Human spliced tran
	27	13.4	67.0	18	21	AAZ74872	Human biallelic ma
	28	13.4	67.0	21	21	AAZ74470	Human biallelic ma
	29	13.4	67.0	43	22	AAF86258	Mutagenic primer c
	30	13.2	66.0	51	21	AAA69522	VEGF antagonist-Fc
c	31	13.2	66.0	51	21	AAA69523	VEGF antagonist-Fc
c	32	13.2	66.0	51	22	AAL28258	Human SNP oligonuc
	33	13.2	66.0	51	24	ABL35755	VEGFantagonist-FC
c	34	13.2	66.0	51	24	ABL35756	VEGFantagonist-FC
	35	13.2	66.0	65	24	ABN28311	Rat spliced transc
	36	13.2	66.0	65	24	ABN29769	Rat spliced transc
	37	13.2	66.0	80	22	ABA76756	Human foetal liver
	38	13.2	66.0	80	22	ABA41250	Probe #19716 for g
	39	13.2	66.0	80	22	AAK51388	Human bone marrow
	40	13.2	66.0	80	22	AAI57465	Probe #26151 used
	41	13.2	66.0	80	24	ABS24942	Human genome-deriv
	42	13.2	66.0	92	21	AAC32117	Human secreted pro
	43	13	65.0	65	24	ABN51670	Mouse spliced tran
	44	12.8	64.0	24	22	AAI66990	Nucleotide sequenc
c	45	12.8	64.0	25	20	AAX90063	Fervidobacterium s

ALIGNMENTS

RESULT 1
AAZ31459
ID AAZ31459 standard; DNA; 20 BP.
XX
AC AAZ31459;
XX
DT 07-FEB-2000 (first entry)
XX
DE Human neuropilin mRNA specific antisense oligo GTI3630.
XX
KW Neuropilin; human; growth; metastasis; tumor; neovascularisation;
KW cancer; papilloma; diabetic retinopathy; antisense; ss.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9955855-A2.
XX
PD 04-NOV-1999.
XX
PF 23-APR-1999; 99WO-CA00324.
XX
PR 23-APR-1998; 98US-0082791.
XX
PA (GENE-) GENESENSE TECHNOLOGIES INC.
XX
PI Wright JA, Young AH, Lee YS;
XX
DR WPI; 2000-023357/02.
XX
PT Antisense oligonucleotides that inhibit neuropilin expression, useful
PT for treating cancer -

```
XX
PS   Claim 4; Page 17; 57pp; English.
XX
CC   Sequences AAZ31431-460 represent antisense oligonucleotides which
CC   inhibit human neuropilin expression. The antisense oligonucleotides can
CC   be used to inhibit the growth or metastasis of a mammalian tumor and
CC   inhibit neovascularisation. The oligonucleotides may be used to treat
CC   various forms of cancers or tumors, such as sarcomas, melanomas,
CC   adenomas, carcinomas of solid tissue, hypoxic tumors, squamous cell
CC   carcinomas of the mouth, throat, larynx and lung, genitourinary cancers
CC   such as cervical and bladder cancer, hematopoietic cancers, colon cancer,
CC   breast cancer, pancreatic cancer, renal cancer, brain cancer, skin
CC   cancer, liver cancer, head and neck cancers, and nervous system cancers,
CC   as well as benign lesions such as papillomas. The methods may be used to
CC   treat neovascularisation disorders such as diabetic retinopathy, and
CC   retinopathy of prematurity and age related macular degeneration.
XX
SQ   Sequence 20 BP; 8 A; 6 C; 2 G; 4 T; 0 other;

  Query Match             100.0%;  Score 20;  DB 21;  Length 20;
  Best Local Similarity   100.0%;  Pred. No. 3.6;
  Matches   20;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 ACCATCCACAAGTTCAAAGT 20
          ||||||||||||||||||||
Db      1 ACCATCCACAAGTTCAAAGT 20


RESULT 2
AAT00985
ID   AAT00985 standard; DNA; 30 BP.
XX
AC   AAT00985;
XX
DT   02-APR-1996  (first entry)
XX
DE   Human papilloma virus HPV-34-specific probe.
XX
KW   Human papilloma virus; HPV-34; cervical carcinoma; screening; PCR;
KW   elongated general primer; nucleic acid amplification; probe; ss.
XX
OS   Synthetic.
XX
PN   WO9522626-A1.
XX
PD   24-AUG-1995.
XX
PF   20-FEB-1995;   95WO-NL00066.
XX
PR   23-SEP-1994;   94EP-0202739.
PR   21-FEB-1994;   94EP-0200432.
XX
PA   (RESE-) STICHTING RESEARCHFONDS PATHOLOGIE.
XX
PI   Meijer CJLM,  Snijders PJF,  Van Den Brule AJC,  Walboomers JMM;
XX
DR   WPI; 1995-302728/39.
XX
PT   Elongated general primer(s) and Human Papilloma Virus-specific
PT   probe(s) - for use in amplification and detection method(s)
PT   providing improved HPV detection in cervical smear(s).
XX
PS   Claim 29; Page 56; 61pp; English.
XX
CC   Elongated general primers GP5+ and GP6+ and their substn. derivs.
CC   are used for amplifying Human Papilloma Virus sequences. Unlike the
CC   known general primers GP5 and GP6, originally selected from the HPV
CC   L1 region on the basis of sequence information of HPV-6, -16, -18,
CC   -31 and -33, the 3'-elongated primers overcome reduced PCR efficiency
CC   and can distinguish HPV status in cytomorphologically normal cervical
CC   smears which previously gave ambiguous or negative results.
CC   The present sequence is that of a probe used for detecting HPV-34
CC   sequences among PCR products amplified by the elongated primers.
XX
SQ   Sequence 30 BP; 14 A; 9 C; 2 G; 5 T; 0 other;

  Query Match             74.0%;  Score 14.8;  DB 16;  Length 30;
  Best Local Similarity   88.9%;  Pred. No. 1.1e+03;
  Matches   16;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 ACCATCCACAAGTTCAAA 18
          || ||||||||||| ||||
Db      4 ACAATCCACAAGTACAAA 21


RESULT 3
AAS14559
ID   AAS14559 standard; DNA; 30 BP.
XX
AC   AAS14559;
XX
DT   18-DEC-2001  (first entry)
XX
DE   HPV oligonucleotide probe for a DNA chip from HPV34.
XX
KW   HPV 34: cervical cancer; endometrial cancer; uterine sarcoma; ss;
KW   probe; genotyping; DNA chip; cytostatic.
XX
OS   Human papillomavirus type 34.
XX
FH   Key             Location/Qualifiers
FT   modified_base   1
FT                   /*tag=  a
FT                   /note= "T is covalently linked to an amine group"
XX
PN   WO200168915-A1.
XX
PD   20-SEP-2001.
XX
PF   26-OCT-2000; 2000WO-KR01213.
XX
PR   15-MAR-2000; 2000KR-0013161.
XX
PA   (BIOM-) BIOMEDLAB CORP.
XX
PI   Park T,  Kim S,  Kim J,  Park M;
XX
DR   WPI; 2001-596914/67.
XX
PT   Human papillomavirus genotyping kit for diagnosing viral infection,
PT   comprises DNA chip with probes, primers for amplifying DNA from sample,
PT   and unit for labeling amplified DNA hybridised with probes of DNA chip
PT   -
XX
PS   Example 1; Page 10; 50pp; English.
XX
CC   The invention relates to a human papillomavirus (HPV) genotyping kit
CC   comprising a DNA chip with probes that have nucleotide sequences
CC   complementary to DNA of HPV, primers for amplifying DNA obtained from
CC   clinical samples by polymerase chain reaction, and a unit for labeling
CC   amplified DNA hybridised with the probes of the DNA chip. The
CC   invention also details preparation of a DNA chip involving preparing 5'
CC   terminal amine-linked DNA probes which have nucleotide sequences
CC   complementary to DNA of HPV, affixing the prepared DNA probes to an
CC   aldehyde-derivatised surface of solid support, and reducing excessive
CC   aldehydes not reacted with amine. The kit is useful for diagnosis of HPV
CC   infection by amplifying DNA obtained from clinical samples by PCR with
CC   primers of the kit to give biotin-containing amplified DNA, applying the
CC   amplified DNA thus obtained to DNA chip of the kit to hybridise the
CC   amplified DNAs with DNA probes of the DNA chip, and detecting DNA bound
CC   on the surface of the DNA chip after labeling amplified DNA hybridised
CC   with the probes with the unit for labeling. The amplification of DNA
CC   obtained from clinical samples is performed by PCR using biotin-16-dUTP.
CC   The kit is useful for identifying types of infecting HPV, and in early
CC   diagnosis, prevention and treatment of cervical cancer,
CC   endometrial cancer and uterine sarcoma. The kit detects HPV infection in
```

```
CC   a simple and accurate manner. The present sequence is an amine-linked
CC   HPV probe from a low risk HPV type.
XX
SQ   Sequence 30 BP; 14 A; 9 C; 2 G; 5 T; 0 other;

  Query Match             74.0%;  Score 14.8;  DB 22;  Length 30;
  Best Local Similarity   88.9%;  Pred. No. 1.1e+03;
  Matches   16;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 ACCATCCACAAGTTCAAA 18
           || |||||||||| ||||
Db       4 ACAATCCACAAGTACAAA 21


RESULT 4
AAC87390/c
ID   AAC87390 standard; DNA; 84 BP.
XX
AC   AAC87390;
XX
DT   09-MAR-2001  (first entry)
XX
DE   Human prostate cancer-predisposing gene CA7 CG04 exon 5.
XX
KW   Human; CA7 CG04 gene; prostate cancer-predisposing gene; exon 5;
KW   chromosome 1 HPC1 region; inherited early onset prostate cancer;
KW   guanosine exchange factor protein; GEF homologue; ras activator;
KW   diagnosis; anticancer drug screening; mutation screening; ds.
XX
OS   Homo sapiens.
XX
PN   WO200069879-A2.
XX
PD   23-NOV-2000.
XX
PF   11-MAY-2000; 2000WO-US12917.
XX
PR   14-MAY-1999;   99US-0134209.
XX
PA   (MYRI-) MYRIAD GENETICS INC.
PA   (HOSP-) HOSPITAL FOR SICK CHILDREN.
XX
PI   Tavtigian SV,  Swedlund B,  Simard J,  Rommens JM;
XX
DR   WPI; 2001-016208/02.
XX
PT   Novel human prostate cancer predisposing gene useful for diagnosis and
PT   prognosis of cancer, especially, prostate cancer and for screening
PT   drugs for cancer therapy -
XX
PS   Claim 7; Page 93; 96pp; English.
XX
CC   The invention relates to a human prostate cancer-predisposing gene,
CC   designated CA7 CG04 (cDNA given in AAC87385), and the CA7 CG04 protein
CC   (AAB48789). The CA7 CG04 protein has homology with guanosine exchange
CC   factor (GEF) proteins, and is therefore thought to function as an
CC   activator of ras. The CA7 CG04 gene is located on chromosome 1 in
CC   the HPC1 region, which is linked with inherited early onset prostate
CC   cancer. The invention also relates to exons 1-19 of the CA7 CG04 gene
CC   (AAC87386-C87404), allelic variants and mutants of the CA7 CG04 gene,
CC   expression vectors and host cells comprising a CA7 CG04 DNA, recombinant
CC   production of the CA7 CG04 protein, anti-CA7 CG04 antibodies, CA7 CG04
CC   primers and probes, and methods of screening for potential anticancer
CC   drugs which can inhibit the ability of a mutant CA7 CG04 to activate
CC   ras. The CA7 CG04 cDNA and protein are useful for screening potential
CC   anticancer drugs and for screening the CA7 CG04 gene for mutations which
CC   are useful in diagnosing a predisposition to cancer, especially prostate
CC   cancer. Sequences AAC87386-C87404 represent exons 1-19 of the human
CC   CA7 CG04 gene.
XX
SQ   Sequence 84 BP; 33 A; 13 C; 16 G; 22 T; 0 other;

  Query Match             74.0%;  Score 14.8;  DB 22;  Length 84;
  Best Local Similarity   88.9%;  Pred. No. 1.2e+03;
  Matches   16;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       2 CCATCCACAAGTTCAAAG 19
           |||||||||| || ||||
Db      21 CCATCCACAACTTGAAAG 4


RESULT 5
ABN29990
ID   ABN29990 standard; DNA; 65 BP.
XX
AC   ABN29990;
XX
DT   15-JUL-2002  (first entry)
XX
DE   Rat spliced transcript detection oligonucleotide SEQ ID NO:2738.
XX
KW   Human; mouse; rat; splice transcript; detection; RNA transcript;
KW   splice variant; transcriptome; oligonucleotide library; ss.
XX
OS   Rattus norvegicus.
XX
PN   WO200210449-A2.
XX
PD   07-FEB-2002.
XX
PF   20-JUL-2001; 2001WO-IB01903.
XX
PR   28-JUL-2000; 2000US-221607P.
PR   02-MAY-2001; 2001US-287724P.
XX
PA   (COMP-) COMPUGEN INC.
XX
PI   Shoshan A,  Wasserman A,  Mintz E,  Mintz L,  Faigler S;
XX
DR   WPI; 2002-257383/30.
XX
PT   New oligonucleotide libraries comprising oligonucleotides which
PT   selectively hybridize to mRNAs transcribed from a transcription unit of
PT   a genome, useful for detecting tissue-, pathology-, and
PT   developmental-specific genes  -
XX
PS   Example 1; SEQ ID 2738; 47pp; English.
XX
CC   The present invention describes oligonucleotide libraries for detecting
CC   messenger RNAs that populate a (sub-)transcriptome, where the
CC   (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC   transcription units that populate a genome. The library comprises
CC   several oligonucleotides, each capable of hybridising selectively to a
CC   set of messenger RNAs transcribed from a given transcription unit of
CC   the genome, which encodes one or more messenger RNA splice variants.
CC   The oligonucleotide libraries are useful for detecting mRNAs from a
CC   biological sample, in expression profiling studies, in qualitatively or
CC   quantitatively characterising the corresponding transcriptome, and in
CC   detecting RNA transcripts and splice variants of human or animal
CC   transcriptomes. The libraries may also be used as specialised mini
CC   libraries to detect transcripts of a sub-transcriptome under a
CC   particular biological or pathological state, and so allowing the
CC   detection of tissue- and pathology-specific genes such as those genes
CC   only expressed in specific tissue under a specific pathological
CC   condition; to detect developmental specific genes; and to detect RNA
CC   transcripts and splice variants of a transcriptome of a patient suffering
CC   from a particular disorder. ABN27253 to ABN59589 represent
CC   oligonucleotide sequences from rats, humans and mice, which are used in
CC   the exemplification of the present invention.
CC   N.B. The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 65 BP; 18 A; 24 C; 9 G; 14 T; 0 other;

  Query Match             72.0%;  Score 14.4;  DB 24;  Length 65;
```

```
  Best Local Similarity   93.8%;  Pred. No. 1.8e+03;
  Matches   15;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       2 CCATCCACAAGTTCAA 17
           |||| ||||||||||
Db      39 CCATACACAAGTTCAA 54


RESULT 6
AAT30466
ID   AAT30466 standard; DNA; 28 BP.
XX
AC   AAT30466;
XX
DT   11-FEB-1997  (first entry)
XX
DE   Primer 5'kappa3 for HNK-20 Vkappa chain coding sequence.
XX
KW   Antibody; HNK-20; variable kappa chain; hybridoma; murine; IgA; mouse;
KW   F glycoprotein; respiratory syncytial virus; RSV; constant region gene;
KW   chimeric antibody; isotype-switched antibody; therapy; infection; human;
KW   pneumonia; bronchiolitis; animal; polymerase chain reaction; primer; PCR;
KW   amplify; ss.
XX
OS   Synthetic.
XX
PN   WO9616974-A1.
XX
PD   06-JUN-1996.
XX
PF   01-DEC-1995;   95WO-US15716.
XX
PR   01-DEC-1994;   94US-0348548.
XX
PA   (ORAV-) ORAVAX INC.
XX
PI   Berdoz J,  Kraehenbuhl J;
XX
DR   WPI; 1996-286826/29.
XX
PT   DNA encoding variable region of antibody HNK-20 - for treating
PT   respiratory syncytial virus infection
XX
PS   Example; Page 34; 75pp; English.
XX
CC   AAT30459-T30545 represent amplification primers for the coding sequences
CC   for the variable chains of an antibody produced by the hybridoma cell
CC   line HNK-20.  AAT30464-T30498 represent amplification primers for the
CC   coding sequence for the variable kappa chain of the HNK-20 antibody.
CC   HNK-20 is a murine hybridoma cell line, that produces IgA specific for
CC   the F glycoprotein of respiratory syncytial virus (RSV).  The variable
CC   chain coding sequences (see AAT30456-T30458) were isolated using primers
CC   specific for the 5' untranslated region of the variable region, and for
CC   the intron downstream of the rearranged J region.  The amplified
CC   sequences can be inserted into vectors containing heterologous (such as
CC   human) constant region genes, for the production of chimeric and
CC   isotype-switched antibodies.  The antibodies are useful in the treatment
CC   and diagnosis of infection by RSV, such as pneumonia and bronchiolitis,
CC   in humans and animals.  By using genomic DNA as a template, variable
CC   region genes can be isolated without producing fragments that have to be
CC   adapted for recombinant antibody expression.  Also, by using the genomic
CC   DNA, no knowledge of the DNA sequence encoding the target variable
CC   region is required.  Chimeric antibodies produced from the encoded
CC   proteins, that contain the constant region of the host being treated, are
CC   less likely to cause adverse immune reactions.
XX
SQ   Sequence 28 BP; 12 A; 5 C; 6 G; 5 T; 0 other;

  Query Match             71.0%;  Score 14.2;  DB 17;  Length 28;
  Best Local Similarity   84.2%;  Pred. No. 2.1e+03;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 ACCATCCACAAGTTCAAAG 19
             | | || ||||||||||||
Db       2 ATCGTCGACAAGTTCAAAG 20


RESULT 7
AAS11817
ID   AAS11817 standard; DNA; 49 BP.
XX
AC   AAS11817;
XX
DT   24-OCT-2001  (first entry)
XX
DE   Mouse surfactant protein A, SPA, oligonucleotide probe #1.
XX
KW   Mouse; surfactant protein A; SPA; Thyroid transcription factor; TTF-1;
KW   lung cancer; thyroid cancer; ds; probe.
XX
OS   Mus sp.
XX
PN   US2001016352-A1.
XX
PD   23-AUG-2001.
XX
PF   26-MAY-1999;   99US-0320337.
XX
PR   17-MAY-1995;   95US-0442809.
PR   18-MAY-1994;   94US-0245356.
XX
PA   (BOHI/) BOHINSKI R J.
PA   (WHIT/) WHITSETT J A.
XX
PI   Bohinski RJ,  Whitsett JA;
XX
DR   WPI; 2001-513959/56.
XX
PT   Oligonucleotide sequences which bind nuclear proteins and surfactants
PT   found in lung cells, useful for detecting cancers that originate in the
PT   lung -
XX
PS   Example 4; Fig 25; 77pp; English.
XX
CC   The invention relates to an oligonucleotide which includes at least 1
CC   nucleic acid sequence which binds to at least 1 nuclear protein found in
CC   lung cells (e.g. the thyroid transcription factor 1, TTF-1, protein).
CC   The oligonucleotide can be expressed in lung cells via a vector and
CC   can be used to target therapeutic agents to kill lung or thyroid cancer
CC   cells. The oligonucleotide can be used to detect or diagnose lung
CC   or thyroid cancer. The oligonucleotides may be designed from the
CC   sequences of, for example, the promoters of lung-specific genes such as
CC   those encoding surfactant proteins. The present sequence is a mouse
CC   surfactant protein A, SPA, probe and is used to identify TTF-1 and HNF-3
CC   binding sites in the SPA promoter in the region -231 to -168.
XX
SQ   Sequence 49 BP; 11 A; 14 C; 9 G; 15 T; 0 other;

  Query Match             71.0%;  Score 14.2;  DB 22;  Length 49;
  Best Local Similarity   84.2%;  Pred. No. 2.2e+03;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 CCATCCACAAGTTCAAAGT 20
           ||| || ||||||| ||||
Db       2 CCACCCTCAAGTTCTAAGT 20


RESULT 8
ABS18070
ID   ABS18070 standard; DNA; 91 BP.
XX
AC   ABS18070;
XX
DT   19-AUG-2002  (first entry)
XX
DE   Human genome-derived single exon probe ORF from lung SEQ ID No 18061.
```

```
XX
KW   Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;
KW   chronic obstructive pulmonary disease; interstitial lung disease;
KW   familial idiopathic pulmonary fibrosis; neurofibromatosis;
KW   tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
KW   Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
KW   pulmonary histiocytosis; lymphangioleiomyomtosis; Karagener syndrome;
KW   pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
KW   primary ciliary dyskinesis; pulmonary hypertension;
KW   hyaline membrane disease; open reading frame; ORF.
XX
OS   Homo sapiens.
XX
PN   WO200186003-A2.
XX
PD   15-NOV-2001.
XX
PF   30-JAN-2001; 2001WO-US00665.
XX
PR   04-FEB-2000; 2000US-180312P.
PR   26-MAY-2000; 2000US-207456P.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-234687P.
PR   27-SEP-2000; 2000US-236359P.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2002-114183/15.
XX
PT   Spatially-addressable set of single exon nucleic acid probes, used to
PT   measure gene expression in human lung samples -
XX
PS   Claim 4; SEQ ID No 18061; 634pp; English.
XX
CC   The invention relates to a spatially-addressable set of single exon
CC   nucleic acid probes for measuring gene expression in a sample derived
CC   from human lung comprising single exon nucleic acid probes having one of
CC   12614 nucleic acid sequences mentioned in the specification, or their
CC   complements or the 12387 open reading frames derived from the 12614
CC   probes. Also included are a microarray comprising the novel set of
CC   probes; the novel set of probes which hybridise at high stringency to a
CC   nucleic acid expressed in the human lung; measuring gene expression in a
CC   sample derived from human lung, comprising (a) contacting the array with
CC   a collection of detectably labeled nucleic acids derived from human lung
CC   mRNA, and (b) measuring the label detectably bound to each probe of
CC   the array; identifying exons in a eukaryotic genome, comprising
CC   (a) algorithmically predicting at least one exon from genomic sequences
CC   of the eukaryote; and (b) detecting specific hybridisation of detectably
CC   labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
CC   having a fragment identical to the predicted exon, the probe is included
CC   in the above mentioned microarray; assigning exons to a single gene,
CC   comprising (a) identifying exons from genomic sequence by the method
CC   above and (b) measuring the expression of each of the exons in several
CC   tissues and/or cell types using hybridisation to a single exon
CC   microarrays having a probe with the exon, where a common pattern of
CC   expression of the exons in the tissues and/or cell types indicates that
CC   the exons should be assigned to a single gene; a peptide comprising one
CC   of 12011 sequences, mentioned in the specification, or encoded by the
CC   probes/open reading frames (ORF). The probes are used for gene
CC   expression analysis, and for identifying exons in a gene, particularly
CC   using human lung derived mRNA and for the study of lung diseases
CC   such as asthma, lung cancer, chronic obstructive pulmonary disease
CC   (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
CC   fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,
CC   Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary
CC   haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomtosis,
CC   pulmonary alveolar proteinosis, Karagener syndrome, fibrocystic
CC   pulmonary dysplasia, primary ciliary dyskinesis, pulmonary hypertension
CC   and hyaline membrane disease. The present sequence is a single exon
CC   probe open reading frame of the invention.
CC   Note: The sequence data for this patent did not form part
CC   of the printed specification, but was obtained in electronic
CC   format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 91 BP; 10 A; 33 C; 7 G; 41 T; 0 other;

  Query Match          71.0%;  Score 14.2;  DB 24;  Length 91;
  Best Local Similarity   84.2%;  Pred. No. 2.4e+03;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 CCATCCACAAGTTCAAAGT 20
           |||||||  |||||||| |
Db      47 CCATCCATCAGTTCAAAAT 65


RESULT 9
ABN93486/c
ID   ABN93486 standard; DNA; 22 BP.
XX
AC   ABN93486;
XX
DT   23-JUL-2002  (first entry)
XX
DE   Human gene GS913018 PCR primer #2.
XX
KW   Human; chromosome 9q31-34; lipoprotein metabolism disorder;
KW   cholesterol transport disorder; PCR; primer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200071710-A2.
XX
PD   30-NOV-2000.
XX
PF   25-MAY-2000; 2000WO-FR01426.
XX
PR   25-MAY-1999;   99FR-0006587.
PR   16-JUN-1999;   99US-0139450.
XX
PA   (AVET ) AVENTIS PHARMA SA.
XX
PI   Denefle P,  Rosier-Montus M,  Arnould-Reguigne I,  Prades C,  Clepet C;
XX
DR   WPI; 2001-025161/03.
XX
PT   New nucleic acid derived from human chromosome 9, used e.g. for
PT   diagnosis and drug screening, derived from genes implicated in
PT   disorders of lipoprotein metabolism  -
XX
PS   Claim 9; Page 224; 269pp; French.
XX
CC   The present invention relates to coding sequences for human genes from
CC   chromosome 9q31-34 (ABN93375-ABN93455). The genes are likely to be
CC   involved in diseases of plasmatic lipoprotein metabolism, e.g. the
CC   reverse transport of cholesterol. The present sequence is a PCR primer
CC   used to illustrate the invention.
XX
SQ   Sequence 22 BP; 3 A; 3 C; 8 G; 8 T; 0 other;

  Query Match          69.0%;  Score 13.8;  DB 23;  Length 22;
  Best Local Similarity   88.2%;  Pred. No. 3.2e+03;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 ACCATCCACAAGTTCAA 17
           || ||||||| ||||||
Db      21 ACAATCCACAGGTTCAA 5


RESULT 10
AAF25277
ID   AAF25277 standard; DNA; 27 BP.
```

```
XX
AC   AAF25277;
XX
DT   30-APR-2001  (first entry)
XX
DE   PCR primer used to amplify DNA encoding human Eag2.
XX
KW   Alpha-subunit; voltage-gated potassium channel; Eag2; ion flux;
KW   central nervous system disorder; migraine; hearing; vision; stroke;
KW   Alzheimer's disease; memory disorder; seizure; psychotic disorder;
KW   PCR primer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200104133-A1.
XX
PD   18-JAN-2001.
XX
PF   12-JUL-2000; 2000WO-US18898.
XX
PR   13-JUL-1999;   99US-0143467.
XX
PA   (ICAG-) ICAGEN INC.
XX
PI   Jegla TJ,  Liu Y;
XX
DR   WPI; 2001-138308/14.
XX
PT   Novel alpha subunit of potassium channel for identifying modulators of
PT   the channel for use in treating disorders involving abnormal ion flux,
PT   e.g. central nervous system disorders -
XX
PS   Claim 7; Page 63; 75pp; English.
XX
CC   PCR primers AAF25272-77 were used to amplify DNA encoding an
CC   alpha-subunit of a voltage-gated potassium channel. The polypeptide is
CC   designated Eag2. The polypeptide is useful for screening activators or
CC   inhibitors of voltage-gated potassium channels that contain an Eag2
CC   subunit. Modulators of voltage-gated channel activity are useful for
CC   treating disorders involving abnormal ion flux, e.g. central nervous
CC   system (CNS) disorders such as migraines, hearing and vision problems,
CC   Alzheimer's disease, learning and memory disorders, seizures, psychotic
CC   disorders and as neuroprotective agents e.g. to prevent stroke. Eag2 is
CC   useful as a reporter molecule to measure changes in potassium
CC   concentration, membrane potential, current flow, ion flux, transcription
CC   signal transduction, receptor-ligand interactions, second messenger
CC   concentration in vitro, in vivo and ex vivo, and also as indicator of
CC   current flow in a particular direction. Detecting Eag2 nucleic acid
CC   and protein expression is useful for diagnosing disease caused by
CC   abnormal ion flux. Eag2 nucleotide and amino acid sequence information
CC   may also be used to construct models of voltage-gated potassium channels
CC   in a computer system.
XX
SQ   Sequence 27 BP; 7 A; 9 C; 3 G; 8 T; 0 other;

  Query Match             69.0%;  Score 13.8;  DB 22;  Length 27;
  Best Local Similarity   88.2%;  Pred. No. 3.2e+03;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       3 CATCCACAAGTTCAAAG 19
           ||||||||  |||||||
Db       9 CATCCACATTTTCAAAG 25


RESULT 11
AAZ38999
ID   AAZ38999 standard; DNA; 30 BP.
XX
AC   AAZ38999;
XX
DT   22-FEB-2000  (first entry)
XX
DE   Human DNA fragmentation factor DFF40 PCR primer SEQ ID NO:10.
XX
KW   Human; DNA fragmentation factor; DFF40; DFF45; apoptosis; gene therapy;
KW   cytostatic; growth; tumour; PCR primer; ss.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
PN   WO9954482-A1.
XX
PD   28-OCT-1999.
XX
PF   16-APR-1998;   98WO-US07895.
XX
PR   16-APR-1998;   98WO-US07895.
XX
PA   (TEXA ) UNIV TEXAS SYSTEM.
XX
PI   Wang X,  Liu X;
XX
DR   WPI; 2000-052702/04.
XX
PT   DNA fragmentation factor DFF40 involved in apoptosis and related
PT   polynucleotide -
XX
PS   Example 1; Page 101; 154pp; English.
XX
CC   The present invention describes a human DNA fragmentation factor,
CC   designated DFF40. Also described are: (1) a method of inducing apoptosis
CC   in a cell comprising providing the cell with DFF40 which results in
CC   apoptosis; (2) a method for inhibiting the growth of a cancer cell
CC   comprising contacting a cancer cell with a DNA fragmentation factor
CC   designated DFF40 under conditions permitting the uptake of the DNA
CC   fragmentation factor by the cell where the presence of the DFF40 into
CC   the cell induces apoptosis; (3) a method for treating cancer comprising:
CC   (a) encoding a DFF40 DNA fragmentation factor; and (b) a promoter active
CC   in the tumour cell, where the promoter is operably linked to the region
CC   encoding the DNA fragmentation factor, under conditions permitting the
CC   uptake of the nucleic acid by the tumour cell; (4) a method of
CC   identifying a modulator of DFF40; and (5) a method of producing a
CC   functional DNA fragmentation factor. An expression construct encoding a
CC   DNA fragmentation factor DFF40 and DFF45 complex is provided to a cell to
CC   induce apoptosis, especially in tumour cells. DFF40 is used to inhibit
CC   the growth of a cancer cell, especially in humans. The present sequence
CC   represents a PCR primer for DFF40, which is used in an example from
CC   the present invention.
XX
SQ   Sequence 30 BP; 13 A; 11 C; 3 G; 3 T; 0 other;

  Query Match             69.0%;  Score 13.8;  DB 21;  Length 30;
  Best Local Similarity   88.2%;  Pred. No. 3.3e+03;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 ACCATCCACAAGTTCAA 17
             |||| ||||||| ||||
Db      13 ACCACCCACAAGCTCAA 29


RESULT 12
AAH74649
ID   AAH74649 standard; DNA; 30 BP.
XX
AC   AAH74649;
XX
DT   15-OCT-2001  (first entry)
XX
DE   PCR primer for cDNA encoding DNA fragmentation factor 40 (DFF40).
XX
KW   Human; DNA fragmentation factor; DFF40; DFF45; apoptosis; DNase;
KW   molecular chaperone; cancer cell; PCR primer; ss.
XX
OS   Homo sapiens.
XX
PN   US2001011078-A1.
```

```
XX
PD   02-AUG-2001.
XX
PF   22-DEC-2000; 2000US-0748451.
XX
PR   16-APR-1998;   98US-0061702.
XX
PA   (TEXA ) UNIV TEXAS SYSTEM.
XX
PI   Wang X,  Liu X;
XX
DR   WPI; 2001-496169/54.
XX
PT   New DNA fragmentation factor polypeptides and polynucleotides, useful
PT   for inhibiting the growth of cancer cells, as well as for inducing
PT   apoptosis of cells -
XX
PS   Example 1; Page 35; 56pp; English.
XX
CC   PCR primers AAH74648-49 were used to amplify cDNA encoding a human DNA
CC   fragmentation factor subunit of 40 kDa, designated DFF40. The
CC   specification also describes DFF45. DFF40 is capable of inducing
CC   apoptosis, and may contain a nuclear localisation fragment. DFF45 acts
CC   as a molecular chaperone to direct the folding of DFF40. Although all
CC   DNase activity is associated with DFF40, DFF activity only occurs once
CC   DFF40 is complexed with DFF45. The DFF polypeptides and polynucleotides
CC   are useful for inhibiting the growth of cancer cells, and for inducing
CC   apoptosis of cells.
XX
SQ   Sequence 30 BP; 13 A; 11 C; 3 G; 3 T; 0 other;

  Query Match             69.0%;  Score 13.8;  DB 22;  Length 30;
  Best Local Similarity   88.2%;  Pred. No. 3.3e+03;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 ACCATCCACAAGTTCAA 17
             |||| ||||||| ||||
Db      13 ACCACCCACAAGCTCAA 29


RESULT 13
AAD02505
ID   AAD02505 standard; DNA; 30 BP.
XX
AC   AAD02505;
XX
DT   24-APR-2001  (first entry)
XX
DE   Human DNA fragmentation factor 40 DNA amplifying forward PCR primer #4.
XX
KW   Human; DNA fragmentation factor; DFF; apoptosis; molecular chaperone;
KW   gene therapy; hyperproliferative disorder; therapy; tumour; restenosis;
KW   psoriasis; angiogenesis; cancer; cytostatic; neoplasia; PCR primer; ss.
XX
OS   Homo sapiens.
XX
PN   US6165737-A.
XX
PD   26-DEC-2000.
XX
PF   16-APR-1998;   98US-0061702.
XX
PR   16-APR-1998;   98US-0061702.
XX
PA   (TEXA ) UNIV TEXAS SYSTEM.
XX
PI   Wang X,  Liu X;
XX
DR   WPI; 2001-090481/10.
XX
PT   Identifying modulator of human DNA fragmentation factor 40, for
PT   treating cancer, involves contacting cell or cell-free composition
PT   comprising DFF40 with candidate substance and comparing apoptosis with
PT   control -
XX
PS   Example 1; Column 63; 52pp; English.
XX
CC   The present sequence is a forward PCR primer used to amplify human DNA
CC   fragmentation factor 40 (DFF40) DNA.
CC   DFF40 is capable of inducing apoptosis. DFF is a heterodimeric protein
CC   comprising 40KDa and 45KDa subunits. DFF45 acts as a molecular chaperone
CC   to facilitate the appropriate folding of DFF40 and acts as an inhibitor
CC   for DFF40. DFF40 and DFF45 are used in gene therapy. The modulators of
CC   human DFF40 activity are useful for inducing apoptosis and for treating
CC   hyperproliferative disorders such as restenosis, psoriasis, metastatic
CC   tumours, angiogenesis and benign and malignant neoplasias. They are also
CC   used for treating cancers of the brain (glioblastoma, astrocytoma,
CC   oligodendroglioma and ependymoma), lung, liver, spleen, kidney, lymph
CC   node, pancreas, small intestine, blood cells, colon, stomach, breast,
CC   endometrium, prostate, testicle, ovary, skin, head and neck, oesophagus,
CC   bone marrow, blood, other tissue and multi-drug resistant cancer.
XX
SQ   Sequence 30 BP; 13 A; 11 C; 3 G; 3 T; 0 other;

  Query Match             69.0%;  Score 13.8;  DB 22;  Length 30;
  Best Local Similarity   88.2%;  Pred. No. 3.3e+03;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 ACCATCCACAAGTTCAA 17
             |||| ||||||| ||||
Db      13 ACCACCCACAAGCTCAA 29


RESULT 14
ABN28293/c
ID   ABN28293 standard; DNA; 65 BP.
XX
AC   ABN28293;
XX
DT   15-JUL-2002  (first entry)
XX
DE   Rat spliced transcript detection oligonucleotide SEQ ID NO:1041.
XX
KW   Human; mouse; rat; splice transcript; detection; RNA transcript;
KW   splice variant; transcriptome; oligonucleotide library; ss.
XX
OS   Rattus norvegicus.
XX
PN   WO200210449-A2.
XX
PD   07-FEB-2002.
XX
PF   20-JUL-2001; 2001WO-IB01903.
XX
PR   28-JUL-2000; 2000US-221607P.
PR   02-MAY-2001; 2001US-287724P.
XX
PA   (COMP-) COMPUGEN INC.
XX
PI   Shoshan A,  Wasserman A,  Mintz E,  Mintz L,  Faigler S;
XX
DR   WPI; 2002-257383/30.
XX
PT   New oligonucleotide libraries comprising oligonucleotides which
PT   selectively hybridize to mRNAs transcribed from a transcription unit of
PT   a genome, useful for detecting tissue-, pathology-, and
PT   developmental-specific genes -
XX
PS   Example 1; SEQ ID 1041; 47pp; English.
XX
CC   The present invention describes oligonucleotide libraries for detecting
CC   messenger RNAs that populate a (sub-)transcriptome, where the
CC   (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC   transcription units that populate a genome. The library comprises
CC   several oligonucleotides, each capable of hybridising selectively to a
CC   set of messenger RNAs transcribed from a given transcription unit of
```

```
CC   the genome, which encodes one or more messenger RNA splice variants.
CC   The oligonucleotide libraries are useful for detecting mRNAs from a
CC   biological sample, in expression profiling studies, in qualitatively or
CC   quantitatively characterising the corresponding transcriptome, and in
CC   detecting RNA transcripts and splice variants of human or animal
CC   transcriptomes. The libraries may also be used as specialised mini
CC   libraries to detect transcripts of a sub-transcriptome under a
CC   particular biological or pathological state, and so allowing the
CC   detection of tissue- and pathology-specific genes such as those genes
CC   only expressed in specific tissue under a specific pathological
CC   condition; to detect developmental specific genes; and to detect RNA
CC   transcripts and splice variants of a transcriptome of a patient suffering
CC   from a particular disorder. ABN27253 to ABN59589 represent
CC   oligonucleotide sequences from rats, humans and mice, which are used in
CC   the exemplification of the present invention.
CC   N.B. The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 65 BP; 12 A; 14 C; 18 G; 21 T; 0 other;

  Query Match             69.0%;  Score 13.8;  DB 24;  Length 65;
  Best Local Similarity   88.2%;  Pred. No. 3.5e+03;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       3 CATCCACAAGTTCAAAG 19
           ||| |||||| ||||||
Db      56 CATACACAAGGTCAAAG 40


RESULT 15
ABN55280
ID   ABN55280 standard; DNA; 65 BP.
XX
AC   ABN55280;
XX
DT   15-JUL-2002  (first entry)
XX
DE   Mouse spliced transcript detection oligonucleotide SEQ ID NO:28028.
XX
KW   Human; mouse; rat; splice transcript; detection; RNA transcript;
KW   splice variant; transcriptome; oligonucleotide library; ss.
XX
OS   Mus musculus.
XX
PN   WO200210449-A2.
XX
PD   07-FEB-2002.
XX
PF   20-JUL-2001; 2001WO-IB01903.
XX
PR   28-JUL-2000; 2000US-221607P.
PR   02-MAY-2001; 2001US-287724P.
XX
PA   (COMP-) COMPUGEN INC.
XX
PI   Shoshan A,  Wasserman A,  Mintz E,  Mintz L,  Faigler S;
XX
DR   WPI; 2002-257383/30.
XX
PT   New oligonucleotide libraries comprising oligonucleotides which
PT   selectively hybridize to mRNAs transcribed from a transcription unit of
PT   a genome, useful for detecting tissue-, pathology-, and
PT   developmental-specific genes  -
XX
PS   Example 1; SEQ ID 28028; 47pp; English.
XX
CC   The present invention describes oligonucleotide libraries for detecting
CC   messenger RNAs that populate a (sub-)transcriptome, where the
CC   (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC   transcription units that populate a genome. The library comprises
CC   several oligonucleotides, each capable of hybridising selectively to a
CC   set of messenger RNAs transcribed from a given transcription unit of
```

```
CC   the genome, which encodes one or more messenger RNA splice variants.
CC   The oligonucleotide libraries are useful for detecting mRNAs from a
CC   biological sample, in expression profiling studies, in qualitatively or
CC   quantitatively characterising the corresponding transcriptome, and in
CC   detecting RNA transcripts and splice variants of human or animal
CC   transcriptomes. The libraries may also be used as specialised mini
CC   libraries to detect transcripts of a sub-transcriptome under a
CC   particular biological or pathological state, and so allowing the
CC   detection of tissue- and pathology-specific genes such as those genes
CC   only expressed in specific tissue under a specific pathological
CC   condition; to detect developmental specific genes; and to detect RNA
CC   transcripts and splice variants of a transcriptome of a patient suffering
CC   from a particular disorder. ABN27253 to ABN59589 represent
CC   oligonucleotide sequences from rats, humans and mice, which are used in
CC   the exemplification of the present invention.
CC   N.B. The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 65 BP; 22 A; 16 C; 14 G; 13 T; 0 other;

  Query Match             69.0%;  Score 13.8;  DB 24;  Length 65;
  Best Local Similarity   88.2%;  Pred. No. 3.5e+03;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       3 CATCCACAAGTTCAAAG 19
           ||||||||| ||||| |
Db      49 CATCCACAACTTCAAGG 65


Search completed: November 23, 2002, 07:04:04
Job time : 98.1 secs
```

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 06:42:25 ; Search time 16.8 Seconds
(without alignments)
450.869 Million cell updates/sec

Title: US-09-296-264-29
Perfect score: 20
Sequence: 1 accatccacaagttcaaagt 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 335578 seqs, 189365133 residues

Total number of hits satisfying chosen parameters: 195614

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	14.2	71.0	91	10	US-09-864-761-30841	Sequence 30841, A
	2	13.8	69.0	30	10	US-09-748-451-10	Sequence 10, Appl
c	3	13.8	69.0	80	10	US-09-864-761-21315	Sequence 21315, A
	4	13.2	66.0	80	10	US-09-864-761-26570	Sequence 26570, A
c	5	12.8	64.0	33	10	US-09-766-095-17	Sequence 17, Appl
	6	12.8	64.0	33	10	US-09-766-095-97	Sequence 97, Appl
c	7	12.8	64.0	33	10	US-09-766-095-134	Sequence 134, App
	8	12.8	64.0	33	10	US-09-766-095-138	Sequence 138, App
c	9	12.8	64.0	42	10	US-09-894-788-12	Sequence 12, Appl
c	10	12.8	64.0	42	10	US-09-894-796-12	Sequence 12, Appl
c	11	12.8	64.0	47	10	US-09-894-788-8	Sequence 8, Appli
c	12	12.8	64.0	47	10	US-09-894-796-8	Sequence 8, Appli
	13	12.8	64.0	52	10	US-09-860-996-29	Sequence 29, Appl
c	14	12.8	64.0	52	10	US-09-894-788-4	Sequence 4, Appli
c	15	12.8	64.0	52	10	US-09-894-796-4	Sequence 4, Appli
	16	12.8	64.0	95	10	US-09-864-761-17113	Sequence 17113, A
	17	12.8	64.0	95	10	US-09-864-761-17149	Sequence 17149, A
	18	12.6	63.0	29	9	US-10-136-139-19	Sequence 19, Appl
	19	12.6	63.0	29	10	US-09-900-062-18	Sequence 18, Appl
	20	12.6	63.0	85	10	US-09-878-574-15184	Sequence 15184, A
	21	12.4	62.0	20	12	US-10-000-864-14	Sequence 14, Appl
	22	12.4	62.0	25	9	US-09-946-807-41	Sequence 41, Appl
	23	12.4	62.0	25	10	US-09-795-668-41	Sequence 41, Appl
	24	12.4	62.0	25	10	US-09-795-686-41	Sequence 41, Appl
c	25	12.4	62.0	80	10	US-09-864-761-16921	Sequence 16921, A
	26	12.2	61.0	21	8	US-08-983-605-309	Sequence 309, App
c	27	12.2	61.0	87	10	US-09-864-761-28254	Sequence 28254, A
	28	12	60.0	20	10	US-09-779-307-35	Sequence 35, Appl
	29	12	60.0	20	10	US-09-779-307-38	Sequence 38, Appl
	30	12	60.0	27	10	US-09-529-063-100	Sequence 100, App
c	31	12	60.0	100	10	US-09-983-965-3593	Sequence 3593, Ap
	32	11.8	59.0	19	10	US-09-969-373-2704	Sequence 2704, Ap
c	33	11.8	59.0	29	10	US-09-745-763-184	Sequence 184, App
c	34	11.8	59.0	42	10	US-09-978-274A-15	Sequence 15, Appl
	35	11.8	59.0	42	10	US-09-978-274A-16	Sequence 16, Appl
c	36	11.8	59.0	81	10	US-09-864-761-18720	Sequence 18720, A
c	37	11.8	59.0	86	10	US-09-864-761-31474	Sequence 31474, A
c	38	11.6	58.0	20	10	US-09-758-987-9	Sequence 9, Appli
	39	11.6	58.0	22	9	US-09-886-156-11	Sequence 11, Appl
c	40	11.6	58.0	34	10	US-09-757-417-48	Sequence 48, Appl
	41	11.6	58.0	40	10	US-09-736-084-34	Sequence 34, Appl
	42	11.6	58.0	41	10	US-09-988-899-61	Sequence 61, Appl
	43	11.6	58.0	41	10	US-09-988-899-62	Sequence 62, Appl
	44	11.6	58.0	41	10	US-09-988-899-63	Sequence 63, Appl
	45	11.6	58.0	51	10	US-09-946-175-28	Sequence 28, Appl

ALIGNMENTS

RESULT 1
US-09-864-761-30841
; Sequence 30841, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30

```
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 30841
;   LENGTH: 91
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AC005030.1
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.57
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.46
;   OTHER INFORMATION: NT HIT: X68216.1, EVALUE 1.90e-01
US-09-864-761-30841

  Query Match             71.0%;  Score 14.2;  DB 10;  Length 91;
  Best Local Similarity   84.2%;  Pred. No. 3.3e+02;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 CCATCCACAAGTTCAAAGT 20
          |||||||  |||||||| |
Db     47 CCATCCATCAGTTCAAAAT 65


RESULT 2
US-09-748-451-10
; Sequence 10, Application US/09748451
; Patent No. US20010011078A1
; GENERAL INFORMATION:
;  APPLICANT: WANG, XIAODONG
;  APPLICANT:  LIU, XUESONG
;  TITLE OF INVENTION: DNA FRAGMENTATION FACTOR INVOLVED IN APOPTOSIS
;  FILE REFERENCE: UTSD:546USD1
;  CURRENT APPLICATION NUMBER: US/09/748,451
;  CURRENT FILING DATE:  2000-12-22
;  PRIOR APPLICATION NUMBER: 09/061,702
;  PRIOR FILING DATE: 1998-04-16
;  NUMBER OF SEQ ID NOS: 21
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
;   LENGTH: 30
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-748-451-10

  Query Match             69.0%;  Score 13.8;  DB 10;  Length 30;
  Best Local Similarity   88.2%;  Pred. No. 4.2e+02;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 ACCATCCACAAGTTCAA 17
          |||| ||||||| ||||
Db     13 ACCACCCACAAGCTCAA 29


RESULT 3
US-09-864-761-21315/c
; Sequence 21315, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 21315
;   LENGTH: 80
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AC012443.1
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.3
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.7
;   OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.4
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.7
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.2
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.4
;   OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.3
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.7
;   OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.5
;   OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.3
;   OTHER INFORMATION: EST_HUMAN HIT: AI290556.1, EVALUE 6.10e-02
US-09-864-761-21315

  Query Match             69.0%;  Score 13.8;  DB 10;  Length 80;
  Best Local Similarity   88.2%;  Pred. No. 5e+02;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      3 CATCCACAAGTTCAAAG 19
          ||| |||||| ||||||
Db     22 CATTCACAAGCTCAAAG 6


RESULT 4
US-09-864-761-26570
; Sequence 26570, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
```

```
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 26570
;   LENGTH: 80
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AL035682.16
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.3
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.2
;   OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.3
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.5
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.4
;   OTHER INFORMATION: NT HIT: AL445064.1, EVALUE 6.40e-01
;   OTHER INFORMATION: EST_HUMAN HIT: BF229975.1, EVALUE 2.40e-01
US-09-864-761-26570

  Query Match             66.0%;  Score 13.2;  DB 10;  Length 80;
  Best Local Similarity   83.3%;  Pred. No. 9.8e+02;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      3 CATCCACAAGTTCAAAGT 20
          ||||||||  | ||||||
Db     34 CATCCACACCTGCAAAGT 51


RESULT 5
US-09-766-095-17/c
; Sequence 17, Application US/09766095
; Patent No. US20020062016A1
;  GENERAL INFORMATION:
;       APPLICANT: Sherrol H. McDonough, Thomas B. Ryder,
;                  Yeasing Yang
;       TITLE OF INVENTION: NUCLEIC ACID AMPLIFICATION
;                           OLIGONUCLEOTIDES AND PROBES
;                           TO HUMAN IMMUNODEFICIENCY VIRUS TYPE 1
;       NUMBER OF SEQUENCES: 139
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Lyon & Lyon
;            STREET: 611 West Sixth Street
;            CITY: Los Angeles
;            STATE: California
;            COUNTRY: USA
;            ZIP: 90017
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
;            COMPUTER: IBM PS/2 Model 50Z or 55SX
;            OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
;            SOFTWARE: WordPerfect (Version 5.0)
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/766,095
;            FILING DATE: 18-Jan-2001
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 09/013,406
;            FILING DATE: 26-JAN-01
;            APPLICATION NUMBER: U.S. Serial No. US20020062016A1 07/550,837
;            FILING DATE: 10-Jul-90
;            APPLICATION NUMBER: U.S. Serial No. US20020062016A1 07/379,501
;            FILING DATE: 11-Jul-89
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Warburg, Richard J.
;            REGISTRATION NUMBER: 32,327
;            REFERENCE/DOCKET NUMBER: 196/189
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (213) 489-1600
;            TELEFAX: (213) 955-0440
;            TELEX: 67-3510
;   INFORMATION FOR SEQ ID NO: 17:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 33
;            TYPE: nucleic acid
;            STRANDEDNESS: single
;            TOPOLOGY: linear
;       SEQUENCE DESCRIPTION: SEQ ID NO: 17:
US-09-766-095-17

  Query Match             64.0%;  Score 12.8;  DB 10;  Length 33;
  Best Local Similarity   87.5%;  Pred. No. 1.3e+03;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      3 CATCCACAAGTTCAAA 18
          |||||||| || |||
Db     33 CATCCACAATTTTAAA 18


RESULT 6
US-09-766-095-97
; Sequence 97, Application US/09766095
; Patent No. US20020062016A1
;  GENERAL INFORMATION:
;       APPLICANT: Sherrol H. McDonough, Thomas B. Ryder,
;                  Yeasing Yang
;       TITLE OF INVENTION: NUCLEIC ACID AMPLIFICATION
;                           OLIGONUCLEOTIDES AND PROBES
;                           TO HUMAN IMMUNODEFICIENCY VIRUS TYPE 1
;       NUMBER OF SEQUENCES: 139
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Lyon & Lyon
;            STREET: 611 West Sixth Street
;            CITY: Los Angeles
;            STATE: California
```

```
;             COUNTRY: USA
;             ZIP: 90017
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
;             COMPUTER: IBM PS/2 Model 50Z or 55SX
;             OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
;             SOFTWARE: WordPerfect (Version 5.0)
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/766,095
;             FILING DATE: 18-Jan-2001
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 09/013,406
;             FILING DATE: 26-JAN-01
;             APPLICATION NUMBER: U.S. Serial No. US20020062016A1 07/550,837
;             FILING DATE: 10-Jul-90
;             APPLICATION NUMBER: U.S. Serial No. US20020062016A1 07/379,501
;             FILING DATE: 11-Jul-89
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Warburg, Richard J.
;             REGISTRATION NUMBER: 32,327
;             REFERENCE/DOCKET NUMBER: 196/189
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (213) 489-1600
;             TELEFAX: (213) 955-0440
;             TELEX: 67-3510
;   INFORMATION FOR SEQ ID NO: 97:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 33
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        SEQUENCE DESCRIPTION: SEQ ID NO: 97:
US-09-766-095-97

  Query Match             64.0%;  Score 12.8;  DB 10;  Length 33;
  Best Local Similarity   87.5%;  Pred. No. 1.3e+03;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       3 CATCCACAAGTTCAAA 18
           ||||||||| || |||
Db       1 CATCCACAATTTTAAA 16


RESULT 7
US-09-766-095-134/c
; Sequence 134, Application US/09766095
; Patent No. US20020062016A1
;   GENERAL INFORMATION:
;        APPLICANT: Sherrol H. McDonough, Thomas B. Ryder,
;                   Yeasing Yang
;        TITLE OF INVENTION: NUCLEIC ACID AMPLIFICATION
;                            OLIGONUCLEOTIDES AND PROBES
;                            TO HUMAN IMMUNODEFICIENCY VIRUS TYPE 1
;        NUMBER OF SEQUENCES: 139
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Lyon & Lyon
;             STREET: 611 West Sixth Street
;             CITY: Los Angeles
;             STATE: California
;             COUNTRY: USA
;             ZIP: 90017
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
;             COMPUTER: IBM PS/2 Model 50Z or 55SX
;             OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
;             SOFTWARE: WordPerfect (Version 5.0)
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/766,095
;             FILING DATE: 18-Jan-2001
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 09/013,406
;             FILING DATE: 26-JAN-01
;             APPLICATION NUMBER: U.S. Serial No. US20020062016A1 07/550,837
;             FILING DATE: 10-Jul-90
;             APPLICATION NUMBER: U.S. Serial No. US20020062016A1 07/379,501
;             FILING DATE: 11-Jul-89
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Warburg, Richard J.
;             REGISTRATION NUMBER: 32,327
;             REFERENCE/DOCKET NUMBER: 196/189
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (213) 489-1600
;             TELEFAX: (213) 955-0440
;             TELEX: 67-3510
;   INFORMATION FOR SEQ ID NO: 134:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 33
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        SEQUENCE DESCRIPTION: SEQ ID NO: 134:
US-09-766-095-134

  Query Match             64.0%;  Score 12.8;  DB 10;  Length 33;
  Best Local Similarity   87.5%;  Pred. No. 1.3e+03;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       3 CATCCACAAGTTCAAA 18
           ||||||||| || |||
Db      33 CATCCACAATTTTAAA 18


RESULT 8
US-09-766-095-138
; Sequence 138, Application US/09766095
; Patent No. US20020062016A1
;   GENERAL INFORMATION:
;        APPLICANT: Sherrol H. McDonough, Thomas B. Ryder,
;                   Yeasing Yang
;        TITLE OF INVENTION: NUCLEIC ACID AMPLIFICATION
;                            OLIGONUCLEOTIDES AND PROBES
;                            TO HUMAN IMMUNODEFICIENCY VIRUS TYPE 1
;        NUMBER OF SEQUENCES: 139
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Lyon & Lyon
;             STREET: 611 West Sixth Street
;             CITY: Los Angeles
;             STATE: California
;             COUNTRY: USA
;             ZIP: 90017
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
;             COMPUTER: IBM PS/2 Model 50Z or 55SX
;             OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
;             SOFTWARE: WordPerfect (Version 5.0)
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/766,095
;             FILING DATE: 18-Jan-2001
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 09/013,406
;             FILING DATE: 26-JAN-01
;             APPLICATION NUMBER: U.S. Serial No. US20020062016A1 07/550,837
;             FILING DATE: 10-Jul-90
;             APPLICATION NUMBER: U.S. Serial No. US20020062016A1 07/379,501
;             FILING DATE: 11-Jul-89
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Warburg, Richard J.
;             REGISTRATION NUMBER: 32,327
;             REFERENCE/DOCKET NUMBER: 196/189
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (213) 489-1600
;             TELEFAX: (213) 955-0440
;             TELEX: 67-3510
```

```
;   INFORMATION FOR SEQ ID NO: 138:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 33
;             TYPE: nucleic acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        SEQUENCE DESCRIPTION: SEQ ID NO: 138:
US-09-766-095-138

  Query Match             64.0%;  Score 12.8;  DB 10;  Length 33;
  Best Local Similarity   68.8%;  Pred. No. 1.3e+03;
  Matches   11;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      3 CATCCACAAGTTCAAA 18
          ||:||||| :: |||
Db      1 CAUCCACAAUUUUAAA 16


RESULT 9
US-09-894-788-12/c
; Sequence 12, Application US/09894788
; Patent No. US20020094527A1
; GENERAL INFORMATION:
;  APPLICANT: Nadeau, James G.
;  APPLICANT:  Hellyer, Tobin J.
;  TITLE OF INVENTION: Probes and Methods for Detection of Nucleic Acids
;  FILE REFERENCE: Universal Reporter
;  CURRENT APPLICATION NUMBER: US/09/894,788
;  CURRENT FILING DATE:  2001-06-28
;  PRIOR APPLICATION NUMBER: 09/590,061
;  PRIOR FILING DATE: 2000-06-08
;  NUMBER OF SEQ ID NOS: 20
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
;   LENGTH: 42
;   TYPE: DNA
;   ORGANISM: Human immunodeficiency virus
US-09-894-788-12

  Query Match             64.0%;  Score 12.8;  DB 10;  Length 42;
  Best Local Similarity   87.5%;  Pred. No. 1.4e+03;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      3 CATCCACAAGTTCAAA 18
          |||||||| || |||
Db     42 CATCCACAATTTTAAA 27


RESULT 10
US-09-894-796-12/c
; Sequence 12, Application US/09894796
; Patent No. US20020102574A1
; GENERAL INFORMATION:
;  APPLICANT: Nadeau, James G.
;  APPLICANT:  Hellyer, Tobin J.
;  TITLE OF INVENTION: Probes and Methods for Detection of Nucleic Acids
;  FILE REFERENCE: Universal Reporter
;  CURRENT APPLICATION NUMBER: US/09/894,796
;  CURRENT FILING DATE:  2001-06-28
;  PRIOR APPLICATION NUMBER: 09/590,061
;  PRIOR FILING DATE: 2000-06-08
;  NUMBER OF SEQ ID NOS: 20
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
;   LENGTH: 42
;   TYPE: DNA
;   ORGANISM: Human immunodeficiency virus
US-09-894-796-12

  Query Match             64.0%;  Score 12.8;  DB 10;  Length 42;
  Best Local Similarity   87.5%;  Pred. No. 1.4e+03;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      3 CATCCACAAGTTCAAA 18
          |||||||| || |||
Db     42 CATCCACAATTTTAAA 27


RESULT 11
US-09-894-788-8/c
; Sequence 8, Application US/09894788
; Patent No. US20020094527A1
; GENERAL INFORMATION:
;  APPLICANT: Nadeau, James G.
;  APPLICANT:  Hellyer, Tobin J.
;  TITLE OF INVENTION: Probes and Methods for Detection of Nucleic Acids
;  FILE REFERENCE: Universal Reporter
;  CURRENT APPLICATION NUMBER: US/09/894,788
;  CURRENT FILING DATE:  2001-06-28
;  PRIOR APPLICATION NUMBER: 09/590,061
;  PRIOR FILING DATE: 2000-06-08
;  NUMBER OF SEQ ID NOS: 20
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
;   LENGTH: 47
;   TYPE: DNA
;   ORGANISM: Human immunodeficiency virus
US-09-894-788-8

  Query Match             64.0%;  Score 12.8;  DB 10;  Length 47;
  Best Local Similarity   87.5%;  Pred. No. 1.4e+03;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      3 CATCCACAAGTTCAAA 18
          |||||||| || |||
Db     47 CATCCACAATTTTAAA 32


RESULT 12
US-09-894-796-8/c
; Sequence 8, Application US/09894796
; Patent No. US20020102574A1
; GENERAL INFORMATION:
;  APPLICANT: Nadeau, James G.
;  APPLICANT:  Hellyer, Tobin J.
;  TITLE OF INVENTION: Probes and Methods for Detection of Nucleic Acids
;  FILE REFERENCE: Universal Reporter
;  CURRENT APPLICATION NUMBER: US/09/894,796
;  CURRENT FILING DATE:  2001-06-28
;  PRIOR APPLICATION NUMBER: 09/590,061
;  PRIOR FILING DATE: 2000-06-08
;  NUMBER OF SEQ ID NOS: 20
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
;   LENGTH: 47
;   TYPE: DNA
;   ORGANISM: Human immunodeficiency virus
US-09-894-796-8

  Query Match             64.0%;  Score 12.8;  DB 10;  Length 47;
  Best Local Similarity   87.5%;  Pred. No. 1.4e+03;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      3 CATCCACAAGTTCAAA 18
          |||||||| || |||
Db     47 CATCCACAATTTTAAA 32


RESULT 13
US-09-860-996-29
; Sequence 29, Application US/09860996
; Patent No. US20020034393A1
; GENERAL INFORMATION:
;  APPLICANT: Mitrophanous, et al
;  TITLE OF INVENTION: VECTOR
;  FILE REFERENCE: 674523-2010
```

; CURRENT APPLICATION NUMBER: US/09/860,996
; CURRENT FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: PCT/GB99/03866
; PRIOR FILING DATE: 1999-11-19
; PRIOR APPLICATION NUMBER: 9825524.3
; PRIOR FILING DATE: 1998-11-20
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 29
; LENGTH: 52
; TYPE: DNA
; ORGANISM: Human Immunodeficiency Virus Type 1
US-09-860-996-29

Query Match 64.0%; Score 12.8; DB 10; Length 52;
Best Local Similarity 87.5%; Pred. No. 1.4e+03;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 CATCCACAAGTTCAAA 18
||||||||| || |||
Db 17 CATCCACAATTTTAAA 32

RESULT 14
US-09-894-788-4/c
; Sequence 4, Application US/09894788
; Patent No. US20020094527A1
; GENERAL INFORMATION:
; APPLICANT: Nadeau, James G.
; APPLICANT: Hellyer, Tobin J.
; TITLE OF INVENTION: Probes and Methods for Detection of Nucleic Acids
; FILE REFERENCE: Universal Reporter
; CURRENT APPLICATION NUMBER: US/09/894,788
; CURRENT FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: 09/590,061
; PRIOR FILING DATE: 2000-06-08
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 52
; TYPE: DNA
; ORGANISM: Human immunodeficiency virus
US-09-894-788-4

Query Match 64.0%; Score 12.8; DB 10; Length 52;
Best Local Similarity 87.5%; Pred. No. 1.4e+03;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 CATCCACAAGTTCAAA 18
||||||||| || |||
Db 52 CATCCACAATTTTAAA 37

RESULT 15
US-09-894-796-4/c
; Sequence 4, Application US/09894796
; Patent No. US20020102574A1
; GENERAL INFORMATION:
; APPLICANT: Nadeau, James G.
; APPLICANT: Hellyer, Tobin J.
; TITLE OF INVENTION: Probes and Methods for Detection of Nucleic Acids
; FILE REFERENCE: Universal Reporter
; CURRENT APPLICATION NUMBER: US/09/894,796
; CURRENT FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: 09/590,061
; PRIOR FILING DATE: 2000-06-08
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 52
; TYPE: DNA
; ORGANISM: Human immunodeficiency virus
US-09-894-796-4

Query Match 64.0%; Score 12.8; DB 10; Length 52;
Best Local Similarity 87.5%; Pred. No. 1.4e+03;
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3 CATCCACAAGTTCAAA 18
||||||||| || |||
Db 52 CATCCACAATTTTAAA 37

Search completed: November 23, 2002, 07:10:44
Job time : 16.8 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 26, 2002, 04:16:10 ; Search time 798.5 Seconds
(without alignments)
405.648 Million cell updates/sec

Title: US-09-296-264-29
Perfect score: 20
Sequence: 1 accatccacaagttcaaagt 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 357874

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	14.8	74.0	67	9	AA625549	AA625549 af72h10.r
	2	14.2	71.0	60	10	BE316566	BE316566 NF057A06L
c	3	14.2	71.0	96	12	BF449377	BF449377 maa42b01.
	4	14.2	71.0	96	17	AZ925323	AZ925323 4910.ez33
	5	13.8	69.0	66	17	BH861238	BH861238 SALK_0551
	6	13.8	69.0	67	17	AZ818261	AZ818261 2M0088K09
	7	13.8	69.0	84	17	AL756367	AL756367 Arabidops
c	8	13.8	69.0	100	17	AZ859074	AZ859074 2M0164H04
c	9	13.6	68.0	81	14	BQ568689	BQ568689 gi112f08.
c	10	13.6	68.0	85	9	AI052301	AI052301 oy93f01.x
	11	13.6	68.0	96	14	T80051	T80051 yd22d01.s1
c	12	13.4	67.0	90	14	N50460	N50460 yy88a07.r1
	13	13.2	66.0	43	17	AZ576107	AZ576107 AST-T29E0
	14	13.2	66.0	47	17	BH813987	BH813987 SALK_0655
	15	13.2	66.0	50	13	BI175138	BI175138 OSTR010D1
	16	13.2	66.0	64	17	BH813872	BH813872 SALK_0654
c	17	13.2	66.0	70	13	BI973022	BI973022 sag86a02.
	18	13.2	66.0	70	14	W85524	W85524 mf58e04.r1
c	19	13.2	66.0	92	14	R29253	R29253 F1-372D 22
c	20	13.2	66.0	98	12	BF971786	BF971786 602240066
	21	13.2	66.0	98	14	BQ640571	BQ640571 he30d07.y
c	22	13	65.0	70	17	AZ766459	AZ766459 1M0564K05
c	23	13	65.0	81	14	R10982	R10982 yf38b10.r1
c	24	13	65.0	84	17	AZ445335	AZ445335 1M0241L05
	25	12.8	64.0	58	14	D25842	D25842 HUMGS04214
	26	12.8	64.0	65	17	BH153111	BH153111 67 Bacter
	27	12.8	64.0	72	14	F32099	F32099 HSPD24466 H
c	28	12.8	64.0	73	13	BI943086	BI943086 sn25f07.y
	29	12.8	64.0	74	9	AA672812	AA672812 vs62b12.r
	30	12.8	64.0	76	9	AA869338	AA869338 vq50g11.r
c	31	12.8	64.0	78	17	BH254261	BH254261 SALK_0162
	32	12.8	64.0	80	14	F34720	F34720 HSPD29981 H
c	33	12.8	64.0	87	10	AV565351	AV565351 AV565351
	34	12.8	64.0	87	17	BH218688	BH218688 1006081C1
	35	12.8	64.0	91	17	BH811585	BH811585 SALK_0591
	36	12.8	64.0	94	9	AA936104	AA936104 oh35c09.s
c	37	12.8	64.0	94	12	BG156488	BG156488 sab09h11.
	38	12.8	64.0	96	14	F32461	F32461 HSPD25222 H
	39	12.8	64.0	96	14	W76234	W76234 zd59c10.r1
c	40	12.8	64.0	98	9	AT005693	AT005693 AT005693
	41	12.8	64.0	100	10	AW454338	AW454338 zehl1308
c	42	12.6	63.0	39	17	BH811645	BH811645 SALK_0595
c	43	12.6	63.0	54	9	AU014366	AU014366 AU014366
c	44	12.6	63.0	57	9	AU010020	AU010020 AU010020
	45	12.6	63.0	73	17	CNS02FGM	AL195007 Tetraodon

ALIGNMENTS

RESULT 1
AA625549
LOCUS AA625549 67 bp mRNA linear EST 02-MAR-1998
DEFINITION af72h10.r1 Soares_NhHMPu_S1 Homo sapiens cDNA clone IMAGE:1047619 5' similar to TR:G1256436 G1256436 APC-BINDING PROTEIN EB2 ;, mRNA sequence.
ACCESSION AA625549
VERSION AA625549.1 GI:2537936
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 67)
AUTHORS Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S., Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B., White,Y., Wylie,T., Waterston,R. and Wilson,R.
TITLE WashU-NCI human EST Project
JOURNAL Unpublished (1997)
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Trace considered overall poor quality

```
          Possible reversed clone: similarity on wrong strand
          Insert Length: 759   Std Error: 0.00
          Seq primer: -28m13 rev2 ET from Amersham
          High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .67
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1047619"
                     /clone_lib="Soares_NhHMPu_S1"
                     /tissue_type="Pooled human melanocyte, fetal heart, and
                     pregnant uterus"
                     /lab_host="DH10B"
                     /note="Organ: mixed (see below); Vector: pT7T3D-Pac
                     (Pharmacia) with a modified polylinker; Site_1: Not I;
                     Site_2: Eco RI; Equal amounts of plasmid DNA from three
                     normalized libraries (melanocyte 2NbHM, pregnant uterus
                     NbHPU, and fetal heart NbHH19W) were mixed, and ss circles
                     were made in vitro. Following HAP purification, this DNA
                     was used as tracer in a subtractive hybridization
                     reaction. The driver was PCR-amplified cDNAs from pools of
                     5,000 clones made from the same 3 libraries. The pools
                     consisted of I.M.A.G.E. clones 260232-265223,
                     340488-345479, and 484488-489479."
BASE COUNT       24 a     15 c     15 g     13 t
ORIGIN

  Query Match             74.0%;  Score 14.8;  DB 9;  Length 67;
  Best Local Similarity   88.9%;  Pred. No. 4.2e+03;
  Matches   16;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       3 CATCCACAAGTTCAAAGT 20
           |||||||||| ||||| ||
Db      43 CATCCACAACTTCAAGGT 60


RESULT 2
BE316566
LOCUS       BE316566                 60 bp    mRNA    linear   EST 21-DEC-2000
DEFINITION  NF057A06LF1F1037 Developing leaf Medicago truncatula cDNA clone
            NF057A06LF 5', mRNA sequence.
ACCESSION   BE316566
VERSION     BE316566.2  GI:11961806
KEYWORDS    EST.
SOURCE      barrel medic.
  ORGANISM  Medicago truncatula
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
            Medicago.
REFERENCE   1  (bases 1 to 60)
  AUTHORS   Torres-Jerez,I., Scott,A.D., Harris,A.R., Gonzales,R.A., Bell,C.J.,
            Flores,H.R., Inman,J.T., Weller,J.W. and May,G.D.
  TITLE     Expressed Sequence Tags from the Samuel Roberts Noble Foundation
            Medicago truncatula leaf library
  JOURNAL   Unpublished (2000)
COMMENT     On Jul 14, 2000 this sequence version replaced gi:9190343.
            Contact: May GD
            Plant Biology Division
            The Samuel Roberts Noble Foundation
            2510 Sam Noble Parkway, Ardmore, OK 73402, USA
            Tel: 580 221 7391
            Fax: 580 221 7380
            Email: gdmay@noble.org
            Medicago Genome Initiative accession: MGI:S:18639
            Insert Length: 808   Std Error: 0.00
            Plate: 057  row: A  column: 06
            Seq primer: TCACACAGGAAACAGCTATGAC.
FEATURES             Location/Qualifiers
     source          1. .60
                     /organism="Medicago truncatula"
                     /db_xref="taxon:3880"
                     /clone="NF057A06LF"
                     /clone_lib="Developing leaf"
                     /tissue_type="leaf"
                     /dev_stage="Pooled developmental"
                     /note="Vector: Lambda Zap; Contains a mixture of very
                     young, developing, mature and senescing leaves."
BASE COUNT       19 a     19 c      8 g     14 t
ORIGIN

  Query Match             71.0%;  Score 14.2;  DB 10;  Length 60;
  Best Local Similarity   84.2%;  Pred. No. 7.9e+03;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 ACCATCCACAAGTTCAAAG 19
           ||| ||||||| ||||| |
Db      24 ACCTTCCACAATTTCAATG 42


RESULT 3
BF449377/c
LOCUS       BF449377                 96 bp    mRNA    linear   EST 01-DEC-2000
DEFINITION  maa42b01.y1 NCI_CGAP_Brn63 Mus musculus cDNA clone IMAGE:3813457
            5', mRNA sequence.
ACCESSION   BF449377
VERSION     BF449377.1  GI:11515546
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 96)
  AUTHORS   NCI/NINDS-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute / National Institute of Neurological
            Disorders and Stroke, Brain Tumor Genome Anatomy Project
            (CGAP/BTGAP), Tumor Gene Index
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Jeffrey E. Green, M.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov

            MGI:1455569
            Seq primer: -40RP from Gibco.
FEATURES             Location/Qualifiers
     source          1. .96
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:3813457"
                     /clone_lib="NCI_CGAP_Brn63"
                     /sex="female"
                     /dev_stage="10 weeks"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
                     Average insert size 1.8 kb.  Library constructed by Life
                     Technologies."
BASE COUNT       43 a     17 c     14 g     22 t
ORIGIN

  Query Match             71.0%;  Score 14.2;  DB 12;  Length 96;
  Best Local Similarity   84.2%;  Pred. No. 9.6e+03;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 CCATCCACAAGTTCAAAGT 20
           |||| |||||||  |||||
Db      25 CCATGCACAAGTAGAAAGT 7


RESULT 4
```

```
AZ925323
LOCUS       AZ925323                96 bp    DNA     linear   GSS 01-APR-2001
DEFINITION  4910.ez33b02.x1 Saccharomyces paradoxus N17 Saccharomyces paradoxus
            genomic clone 4910.ez33b02.x1, DNA sequence.
ACCESSION   AZ925323
VERSION     AZ925323.1  GI:13496223
KEYWORDS    GSS.
SOURCE      Saccharomyces paradoxus.
  ORGANISM  Saccharomyces paradoxus
            Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
            Saccharomycetales; Saccharomycetaceae; Saccharomyces.
REFERENCE   1  (bases 1 to 96)
  AUTHORS   Cliften,P.F., Hillier,L.W., Fulton,L., Graves,T., Miner,T., Gish
            ,W.R., Waterston,R.H. and Johnston,M.
  TITLE     Surveying Saccharomyces genomes to identify functional elements by
            comparative DNA sequence analysis
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Johnston M
            Department of Genetics
            Washington University Medical School
            Box 8232, 4566 Scott Ave., St. Louis, MO 63110, USA
            Tel: 314 362 2735
            Fax: 314 362 7855
            Email: mj@genetics.wustl.edu
            Class: random plasmid subclone.
FEATURES             Location/Qualifiers
     source          1. .96
                     /organism="Saccharomyces paradoxus"
                     /strain="N17"
                     /db_xref="taxon:27291"
                     /clone="4910.ez33b02.x1"
                     /clone_lib="Saccharomyces paradoxus N17"
                     /note="Random genomic sequence "
BASE COUNT       36 a     22 c     15 g     23 t
ORIGIN

  Query Match            71.0%;  Score 14.2;  DB 17;  Length 96;
  Best Local Similarity  84.2%;  Pred. No. 9.6e+03;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 CCATCCACAAGTTCAAAGT 20
          |||||||| | ||||||  |
Db     49 CCATCCACTACTTCAAAAT 67


RESULT 5
BH861238
LOCUS       BH861238                66 bp    DNA     linear   GSS 05-AUG-2002
DEFINITION  SALK_055165 Arabidopsis thaliana TDNA insertion lines Arabidopsis
            thaliana genomic clone SALK_055165, DNA sequence.
ACCESSION   BH861238
VERSION     BH861238.1  GI:22096564
KEYWORDS    GSS.
SOURCE      thale cress.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 66)
  AUTHORS   Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R., Gadrinab
            ,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L., Shinn,P.
            , Zimmerman,J. and Ecker,J.R.
  TITLE     A Sequence-Indexed Library of Insertion Mutations in the
            Arabidopsis Genome
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Joseph R. Ecker
            Salk Institute Genomic Analysis Laboratory (SIGnAL)
            The Salk Institute for Biological Studies
            10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
            Tel: 858 453 4100 x1752
            Fax: 858 558 6379
            Email: ecker@salk.edu
            This is single pass sequence recovered from the left border of
            TDNA.
            Class: TDNA tagged.
FEATURES             Location/Qualifiers
     source          1. .66
                     /organism="Arabidopsis thaliana"
                     /strain="Columbia 0"
                     /db_xref="taxon:3702"
                     /clone="SALK_055165"
                     /clone_lib="Arabidopsis thaliana TDNA insertion lines"
                     /note="PCR was performed on Arabidopsis thaliana lines
                     each of which contains  one or more TDNA insertion
                     elements. The resultant fragment for each  line was
                     directly sequenced to determine the genomic sequence at
                     the  site of insertion. Details of the protocols used can
                     be found at  http://signal.salk.edu/tdna_protocols.html"
BASE COUNT       20 a     10 c     16 g     20 t
ORIGIN

  Query Match            69.0%;  Score 13.8;  DB 17;  Length 66;
  Best Local Similarity  88.2%;  Pred. No. 1.3e+04;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      1 ACCATCCACAAGTTCAA 17
          |||||||||||||  ||
Db     28 ACCATCCACAAGTGGAA 44


RESULT 6
AZ818261
LOCUS       AZ818261                67 bp    DNA     linear   GSS 20-FEB-2001
DEFINITION  2M0088K09F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC2M0088K09 F, DNA sequence.
ACCESSION   AZ818261
VERSION     AZ818261.1  GI:12988169
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 67)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0088  row: K  column: 09
            Seq primer: CGTTGTAAAACGACGGCCAGT
            Class: plasmid ends
            High quality sequence stop: 67.
FEATURES             Location/Qualifiers
     source          1. .67
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC2M0088K09"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
```

was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptored DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptored mouse DNA was annealed to adaptored vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 22 a 22 c 9 g 14 t
ORIGIN

Query Match 69.0%; Score 13.8; DB 17; Length 67;
Best Local Similarity 88.2%; Pred. No. 1.3e+04;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ACCATCCACAAGTTCAA 17
||||||||||| | |||
Db 23 ACCATCCACAATTCCAA 39

RESULT 7
AL756367
LOCUS AL756367 84 bp DNA linear GSS 17-JUN-2002
DEFINITION Arabidopsis thaliana T-DNA flanking sequence GK-108D06-012313, genomic survey sequence.
ACCESSION AL756367
VERSION AL756367.1 GI:21488865
KEYWORDS GSS.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1
AUTHORS Strizhov,N., Li,Y., Rosso,M., Viehoever,P., Dekker,K., Saedler,H. and Weisshaar,B.
TITLE A pipeline for automated high-throughput generation of FSTs (flanking sequence tags) from Arabidopsis thaliana T-DNA transformed lines
JOURNAL Unpublished
REFERENCE 2
AUTHORS Rosso,M., Strizhov,N., Li,Y., Reiss,B., Dekker,K. and Weisshaar,B.
TITLE A new Arabidopsis thaliana T-DNA mutagenised population (GABI-Kat) for flanking sequence tag based reverse genetics
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 84)
AUTHORS Rosso,M., Li,Y., Strizhov,N. and Weisshaar,B.
TITLE Direct Submission
JOURNAL Submitted (17-JUN-2002) Weisshaar B., Max-Planck-Institut fuer Zuechtungsforschung, Carl-von-Linne-Weg 10, Koeln, 50829, Germany
COMMENT This sequence is recovered from the left border of the T-DNA. It indicates an insertion close to or within gene At5g10340. The sequences are generated at the MPI for Plant Breeding Research in the context of the GABI-Kat project. GABI-Kat is part of the German Plant Genomics program designated 'GABI'. Information on line availability can be found at:
http://www.mpiz-koeln.mpg.de/GABI-Kat/.
FEATURES Location/Qualifiers
source 1. .84
/organism="Arabidopsis thaliana"
/strain="Columbia 0"
/db_xref="taxon:3702"
/clone="GK-108D06-012313"
/clone_lib="Arabidopsis thaliana T-DNA insertion lines"
/note="PCR was performed on DNA from Arabidopsis thaliana plants (T1) which were transformed with the T-DNA from vector pAC161. The lines contain one or more T-DNA insertions. The DNA fragment(s) resulting from the PCR were directly sequenced to determine the genomic sequence flanking the insertion. Sequences displaying significant similarity to the A. thalaina nuclear genome sequence were processed for submission. T-DNA derived sequences were removed"

BASE COUNT 26 a 19 c 17 g 22 t
ORIGIN

Query Match 69.0%; Score 13.8; DB 17; Length 84;
Best Local Similarity 88.2%; Pred. No. 1.4e+04;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ACCATCCACAAGTTCAA 17
||| ||| |||||||||
Db 27 ACCCTCCTCAAGTTCAA 43

RESULT 8
AZ859074/c
LOCUS AZ859074 100 bp DNA linear GSS 21-FEB-2001
DEFINITION 2M0164H04R Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC2M0164H04 R, DNA sequence.
ACCESSION AZ859074
VERSION AZ859074.1 GI:13053008
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 100)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D.,Weiss,R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0164 row: H column: 04
Seq primer: CACACAGGAAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 100.
FEATURES Location/Qualifiers
source 1. .100
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC2M0164H04"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptored DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative

of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptored mouse DNA was annealed to adaptored vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 21 a 23 c 24 g 32 t
ORIGIN

Query Match 69.0%; Score 13.8; DB 17; Length 100;
Best Local Similarity 88.2%; Pred. No. 1.5e+04;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy      3 CATCCACAAGTTCAAAG 19
          ||||||||||| || ||
Db     93 CATCCACAAGTACAGAG 77
```

RESULT 9
BQ568689/c
LOCUS BQ568689 81 bp mRNA linear EST 19-JUN-2002
DEFINITION gi112f08.y2 Mouse Organ of Corti cDNA pBluescript Mus musculus cDNA clone gi112f08 5', mRNA sequence.
ACCESSION BQ568689
VERSION BQ568689.1 GI:21472006
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 81)
AUTHORS Kachar,B.
TITLE EST analysis of gene expression in the mouse Organ of Corti at the onset of hearing
JOURNAL Unpublished (2002)
COMMENT Contact: Kachar,B.
Structural Cell Biology
National Institute of Deafness and other Communication Disorders
50/4249 South Drive, NIH, Bethesda, MD 20892-8027, USA
Tel: 301-402-1599
Fax: 301-402-1765
Email: kacharb@nidcd.nih.gov
Plate: 112 row: f column: 08
Seq primer: M13RP1 reverse primer (ABI).
FEATURES Location/Qualifiers
source 1. .81
/organism="Mus musculus"
/strain="BALB/c"
/db_xref="taxon:10090"
/clone="gi112f08"
/clone_lib="Mouse Organ of Corti cDNA pBluescript"
/sex="male and female"
/dev_stage="Post natal day 5 to 13"
/note="Organ: Organ of Corti; Vector: pBluescript; The organ of Corti (OC) was fine dissected from a total of 386 OC as follows: 102 samples from post-natal (P) day 5; 72 from P6; 60 from P7; 46 from P8; 18 from P9; 20 from P10; 14 from P12 and 24 from P13. After killing animals by cervical dislocation followed by decapitation, the bulla was removed and opened in Leibowitz medium. The bony capsule of the cochlea was chipped away, stria vascularis and spiral ligament were removed and the sensory epithelium was carefully dissected out of the modiolus. Total RNA was extracted using the micro Fasttrack kit (catalog # K1593-02; Invitrogen, Carlsbad, CA), according to manufacturer's instructions. Reverse transcription and library construction were carried out with the Uni-Zap XR vector kit (catalog # 237211, Stratagene) and Uni-Zap XR Gigapack III Gold Cloning kit (catalog # 237612), both from Stratagene (La Jolla, CA, USA), according to manufacturer's instructions. Briefly: 1.5 ug mRNA was reverse transcribed using a hybrid oligo(dT) linker-primer that contains an Xho I site. First strand synthesis was primed with the linker- primer and transcribed using Moloney murine leukemia virus reverse transcriptase (MMLV-RT) and 5-methyl dCTP. The second strand was synthesized with DNA polymerase and RNase H. Complementary DNA was blunt ended with Pfu DNA polymerase, ligated with EcoR I adapters in the presence of ligase and digested with Xho I. The cDNA was sequentially size fractionated over Pharmacia Size Sep400 (Pharmacia, Uppsala, Sweden) and Clontech Chroma Spin-1000 (Clontech, Palo Alto, CA) columns to enrich for cDNAs greater than 400bp and 1000 bp , respectively. The cDNA was then directionally ligated to the Uni-ZAP XR vector, which had been predigested with EcoR I and Xho I. The phagemid was packaged with Gigapak III Gold and, upon titration on XL1 Blue MRF' cells, the yield of the phage library was estimated to be 11,100,000 recombinants. Stratagene's ExAssist Interference resistance helper phage (catalogue # 211203) was adopted to rescue plasmid DNA from the phages. Upon plating of the rescued library, individual cDNA clones were selected and grown in 96-well, 2 ml growth plate. Plasmid DNA was purified from 200 ul of saturated culture with the Concert96(TM) plasmid purification kit (Invitrogen, Carlsbad, CA) as instructed by the manufacturer. ESTs from the 5' end of the cDNA clones were generated with the universal M13 reverse primer (CAGGAAACAGCTATGACC) and 25% strength BigDye terminator sequencing chemistry (Applied Biosystems, Foster City, CA). Sequencing reactions were performed on MJ Tetrad thermal cyclers (MJ Research, Waltham, MA), and analyzed on 3700 automated capillary sequencers using POP5 polymer (Applied Biosystems, Foster City, CA). The frequency distribution of the library is as follows: 72% of genes have 1 copy; 14.3% 2; 12% 3-10; 1.4% 11-50 and 0.1% 51-150. As to gene function, 45% of genes are present in GenBank and have know function; 23% have hits in GenBank, but do not have assigned function; 12% are uncharacterized ESTs and 20% are unidentified."

BASE COUNT 25 a 15 c 14 g 27 t
ORIGIN

Query Match 68.0%; Score 13.6; DB 14; Length 81;
Best Local Similarity 80.0%; Pred. No. 1.7e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
Qy      1 ACCATCCACAAGTTCAAAGT 20
          ||||| | |||| || ||||
Db     31 ACCATGCCCAAGGTCCAAGT 12
```

RESULT 10
AI052301/c
LOCUS AI052301 85 bp mRNA linear EST 29-SEP-1998
DEFINITION oy93f01.x1 Soares_fetal_liver_spleen_1NFLS_S1 Homo sapiens cDNA clone IMAGE:1673401 3' similar to TR:Q14692 Q14692 KIAA0187 PROTEIN. ;, mRNA sequence.
ACCESSION AI052301
VERSION AI052301.1 GI:3308292
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 85)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Trace considered overall poor quality
Insert Length: 1011 Std Error: 0.00

```
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 1.
FEATURES              Location/Qualifiers
     source           1. .85
                      /organism="Homo sapiens"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:1673401"
                      /clone_lib="Soares_fetal_liver_spleen_1NFLS_S1"
                      /sex="male"
                      /dev_stage="20 week-post conception fetus"
                      /lab_host="DH10B (ampicillin resistant)"
                      /note="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia)
                      with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
                      This is a subtracted version of the original Soares fetal
                      liver spleen 1NFLS library.   1st strand cDNA was primed
                      with a Pac I - oligo(dT) primer [5'
                      AACTGGAAGAATTAATTAAAGATCTTTTTTTTTTTTTTTTT 3'],
                      double-stranded cDNA was ligated to Eco RI adaptors
                      (Pharmacia), digested with Pac I and cloned into the Pac I
                      and Eco RI sites of the modified pT7T3 vector.  Library
                      went through one round of normalization. Library
                      constructed by Bento Soares and M.Fatima Bonaldo."
BASE COUNT       15 a     29 c     17 g     24 t
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 9;  Length 85;
  Best Local Similarity   80.0%;  Pred. No. 1.8e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 ACCATCCACAAGTTCAAAGT 20
          |||||||| |||| ||  ||
Db     43 ACCATCCAGAAGTCCAGTGT 24


RESULT 11
T80051
LOCUS       T80051                   96 bp    mRNA    linear   EST 15-MAR-1995
DEFINITION  yd22d01.s1 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone
            IMAGE:108961 3', mRNA sequence.
ACCESSION   T80051
VERSION     T80051.1  GI:698560
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 96)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman
            ,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J.,
            Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston
            ,R., Williamson,A., Wohldmann,P. and Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Insert Size: 743
            High quality sequence stops: 79 Source: IMAGE Consortium, LLNL This
            clone is available royalty-free through LLNL ; contact the IMAGE
            Consortium (info@image.llnl.gov) for further information.
            Insert Length: 743   Std Error: 0.00
            Seq primer: -21m13
            High quality sequence stop: 79.
FEATURES              Location/Qualifiers
     source           1. .96
                      /organism="Homo sapiens"
                      /db_xref="GDB:464578"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:108961"
                      /clone_lib="Soares fetal liver spleen 1NFLS"
                      /sex="male"
                      /dev_stage="20 week-post conception fetus"
                      /lab_host="DH10B (ampicillin resistant)"
                      /note="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia)
                      with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
                      1st strand cDNA was primed with a Pac I - oligo(dT) primer
                      [5' AACTGGAAGAATTAATTAAAGATCTTTTTTTTTTTTTTTTT 3'],
                      double-stranded cDNA was ligated to Eco RI adaptors
                      (Pharmacia), digested with Pac I and cloned into the Pac I
                      and Eco RI sites of the modified pT7T3 vector.  Library
                      went through one round of normalization. Library
                      constructed by Bento Soares and M.Fatima Bonaldo."
BASE COUNT       28 a     20 c     12 g     33 t      3 others
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 14;  Length 96;
  Best Local Similarity   80.0%;  Pred. No. 1.9e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 ACCATCCACAAGTTCAAAGT 20
          | | ||||||||  ||||||
Db     29 ATCCTCCACAAGCACAAAGT 48


RESULT 12
N50460/c
LOCUS       N50460                   90 bp    mRNA    linear   EST 14-FEB-1996
DEFINITION  yy88a07.r1 Soares_multiple_sclerosis_2NbHMSP Homo sapiens cDNA
            clone IMAGE:280596 5', mRNA sequence.
ACCESSION   N50460
VERSION     N50460.1  GI:1191626
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 90)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman
            ,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J.,
            Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston
            ,R., Williamson,A., Wohldmann,P. and Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: T7
            High quality sequence stop: 79.
FEATURES              Location/Qualifiers
     source           1. .90
                      /organism="Homo sapiens"
                      /db_xref="GDB:3898972"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:280596"
                      /clone_lib="Soares_multiple_sclerosis_2NbHMSP"
                      /sex="male"
                      /tissue_type="multiple sclerosis lesions"
                      /dev_stage="Age 46"
                      /lab_host="DH10B (ampicillin resistant)"
                      /note="Vector: pT7T3D (Pharmacia) with a modified
                      polylinker V_TYPE: phagemid; Site_1: Not I; Site_2: Eco RI
                      ; 1st strand cDNA was primed with a Not I - oligo(dT)
                      primer [5'
                      TGTTACCAATCTGAAGTGGGAGCGGCCGCATTTTTTTTTTTTTTTTTT 3'],
                      double-stranded cDNA was size selected, ligated to Eco RI
                      adapters (Pharmacia), digested with Not I and cloned into
                      the Not I and Eco RI sites of a modified pT7T3 vector
                      (Pharmacia). Library went through one round of
```

```
                    normalization to a Cot = 5. Library constructed by Bento
                    Soares and M.Fatima Bonaldo. RNA from 4 multiple sclerosis
                    lesions from one patient was kindly provided by Dr. Kevin
                    G. Becker (NINDS/NIH). "
BASE COUNT       23 a     24 c     24 g     19 t
ORIGIN

  Query Match             67.0%;  Score 13.4;  DB 14;  Length 90;
  Best Local Similarity   93.3%;  Pred. No. 2.3e+04;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      6 CCACAAGTTCAAAGT 20
          ||||||||||| |||
Db     68 CCACAAGTTCACAGT 54


RESULT 13
AZ576107
LOCUS       AZ576107                  43 bp    DNA     linear   GSS 06-DEC-2000
DEFINITION  AST-T29E0055 Genetrap T47D Human Breast Carcinoma Library Homo
            sapiens genomic 5', DNA sequence.
ACCESSION   AZ576107
VERSION     AZ576107.1  GI:11562418
KEYWORDS    GSS.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 43)
  AUTHORS   Henkel,G., Liyanage,M., Pratt,E., Huang,D., Riley,M., Bernardino,A.
            , Durick,K. and Pollok,B.
  TITLE     Exon-trap tags from a T47D GenomeScreen(TM) Library
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Greg Henkel
            Gene Expression
            Aurora Biosciences Corp.
            11010 Torreyana Road, San Diego, CA 92121, USA
            Tel: 8584048436
            Fax: 8584046719
            Email: henkelg@aurorabio.com
            Pools of cells were isolated from a GenomeScreen(TM) library. The
            library of cells was generated by retroviral integration of a gene
            tagging element consisting of: 1) A promoterless beta-lactamase
            proceeded by a splice acceptor as a reporter for gene expression;
            2) A promoter driving neomycin resistance followed by a splice
            donor to trap downstream exons. 3' RACE from neomycin gene was
            performed using total RNA from isolated pools. Output was shotgun
            cloned in pAmp-1 and used to transform DH5-alpha competent
            bacteria. 5' ends of reported sequences were immediately preceded
            by splice donor from the trapping construct.
            Class: exon-trapped.
FEATURES             Location/Qualifiers
     source          1. .43
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone_lib="Genetrap T47D Human Breast Carcinoma Library"
                     /tissue_type="Carcinoma"
                     /cell_type="Epithelial"
                     /cell_line="T47D"
                     /note="Organ: Breast; Vector: pAmp-1; 3' RACE of total RNA
                     from genetrap pools; shotgun clone in pAmp-1 and  used to
                     transform DH5-alpha competent bacteria."
BASE COUNT       20 a     10 c      7 g      5 t      1 others
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 17;  Length 43;
  Best Local Similarity   83.3%;  Pred. No. 2.1e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 ACCATCCACAAGTTCAAA 18
          | ||| | ||||||||||
Db      3 ATCATGCCCAAGTTCAAA 20


RESULT 14
BH813987
LOCUS       BH813987                  47 bp    DNA     linear   GSS 02-MAY-2002
DEFINITION  SALK_065563 Arabidopsis thaliana TDNA insertion lines Arabidopsis
            thaliana genomic clone SALK_065563, DNA sequence.
ACCESSION   BH813987
VERSION     BH813987.1  GI:20393095
KEYWORDS    GSS.
SOURCE      thale cress.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 47)
  AUTHORS   Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R., Gadrinab
            ,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L., Shinn,P.
            , Zimmerman,J. and Ecker,J.R.
  TITLE     A Sequence-Indexed Library of Insertion Mutations in the
            Arabidopsis Genome
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Joseph R. Ecker
            Salk Institute Genomic Analysis Laboratory (SIGnAL)
            The Salk Institute for Biological Studies
            10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
            Tel: 858 453 4100 x1752
            Fax: 858 558 6379
            Email: ecker@salk.edu
            This is single pass sequence recovered from the left border of
            TDNA. This sequence lies within 300 bases of the 5' end of
            At2g18790.
            Class: TDNA tagged.
FEATURES             Location/Qualifiers
     source          1. .47
                     /organism="Arabidopsis thaliana"
                     /strain="Columbia 0"
                     /db_xref="taxon:3702"
                     /clone="SALK_065563"
                     /clone_lib="Arabidopsis thaliana TDNA insertion lines"
                     /note="PCR was performed on Arabidopsis thaliana lines
                     each of which contains  one or more TDNA insertion
                     elements. The resultant fragment for each  line was
                     directly sequenced to determine the genomic sequence at
                     the  site of insertion. Details of the protocols used can
                     be found at  http://signal.salk.edu/tdna_protocols.html"
BASE COUNT       14 a     10 c      5 g     18 t
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 17;  Length 47;
  Best Local Similarity   83.3%;  Pred. No. 2.2e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      3 CATCCACAAGTTCAAAGT 20
          | ||| |||||||||| |
Db      3 CTTCCTCAAGTTCAAAAT 20


RESULT 15
BI175138
LOCUS       BI175138                  50 bp    mRNA    linear   EST 09-JUL-2001
DEFINITION  OSTR010D1_1 AD-wrmcDNA Caenorhabditis elegans cDNA similar to
            K09C8.2, mRNA sequence.
ACCESSION   BI175138
VERSION     BI175138.1  GI:14640941
KEYWORDS    EST.
SOURCE      Caenorhabditis elegans.
  ORGANISM  Caenorhabditis elegans
            Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea
            ; Rhabditidae; Peloderinae; Caenorhabditis.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Reboul,J., Vaglio,P., Tzellas,N., Thierry-Mieg,N., Moore,T.,
            Jackson,C., Shin-i,T., Kohara,Y., Thierry-Mieg,D., Thierry-Mieg,J.,
            Lee,H., Hitti,J., Doucette-Stamm,L., Hartley,J.L., Temple,G.F.,
```

```
            Brasch,M.A., Vandenhaute,J., Lamesch,P.E., Hill,D.E. and Vidal,M.
  TITLE     Open-reading-frame sequence tags (OSTs) support the existence of at
            least 17,300 genes in C. elegans
  JOURNAL   Nat. Genet. 27 (3), 332-336 (2001)
  MEDLINE   21135099
COMMENT     Contact: Reboul J, Vaglio P
            Marc Vidal Laboratory
            Dana Farber Cancer Institute
            44 Binney Street, Boston, MA 02115, USA
            Tel: 617 632 5180
            Fax: 617 632 2425
            Email: Jerome_Reboul@dfci.harvard.edu
            Sequence tag of Gateway entry clones. The primers used were
            designed on the predicted protein encoding ORF. C. elegans ORFeome
            cloning project : Contact jerome_reboul@dfci.harvard.edu or
            philippe_vaglio@dfci.harvard.edu
            POLYA=No.
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="Caenorhabditis elegans"
                     /strain="N2"
                     /db_xref="taxon:6239"
                     /clone_lib="AD-wrmcDNA"
                     /sex="Hermaphrodite and male"
                     /tissue_type="whole animal"
                     /dev_stage="mixed stage"
                     /note="The AD-wrmcDNA library was generated with poly(A)+
                     RNA isolated from both hermaphrodite and male N2 worms of
                     all larval stages, embryos, adults and dauers and the
                     subsequent generation of cDNAs by poly(A) priming. The
                     cDNAs were cloned into pPC86"
BASE COUNT       20 a     18 c      7 g      5 t
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 13;  Length 50;
  Best Local Similarity   83.3%;  Pred. No. 2.2e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 ACCATCCACAAGTTCAAA 18
           ||||  |||| |||||||
Db      22 ACCAGTCACACGTTCAAA 39


Search completed: November 26, 2002, 17:58:30
Job time : 806.5 secs
```

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 3, 2002, 18:14:22 ; Search time 351.3 Seconds
(without alignments)
1656.863 Million cell updates/sec

Title: US-09-296-264-30
Perfect score: 20
Sequence: 1 accacagggctcaccaggcg 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 995600

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
c 1	14.2	71.0	23	6	AX350161	AX350161 Sequence
c 2	14.2	71.0	97	6	I91485	I91485 Sequence 19
3	13.8	69.0	78	10	MUSTCDXBA	M37589 Mouse T cel
4	13.8	69.0	78	10	MUSTCDXBB	M37590 Mouse T cel
5	13.8	69.0	78	10	MUSTCAYA	M37591 Mouse T cel
6	13.8	69.0	78	10	MUSTCAYB	M37592 Mouse T cel
7	13.8	69.0	97	9	HUM13COL15	M68981 Human alpha
c 8	13.6	68.0	51	6	AX204305	AX204305 Sequence
c 9	13.6	68.0	70	14	REO2L3A	J02305 Reovirus se
10	13.4	67.0	21	6	AX145934	AX145934 Sequence
11	13.2	66.0	20	6	A48205	A48205 Sequence 4
12	13.2	66.0	20	6	A57298	A57298 Sequence 4
13	13.2	66.0	20	6	AR076724	AR076724 Sequence
14	13.2	66.0	20	6	AR076738	AR076738 Sequence
15	13.2	66.0	20	6	AR182781	AR182781 Sequence
16	13.2	66.0	20	6	AR182795	AR182795 Sequence
c 17	13.2	66.0	37	6	I33378	I33378 Sequence 1
18	13.2	66.0	38	6	AR046388	AR046388 Sequence
19	13.2	66.0	38	6	I53440	I53440 Sequence 11
c 20	13.2	66.0	50	6	AR074563	AR074563 Sequence
21	13.2	66.0	50	6	AR088044	AR088044 Sequence
c 22	13.2	66.0	50	6	AR157443	AR157443 Sequence
23	13.2	66.0	50	6	E09873	E09873 RT primer.
c 24	13.2	66.0	75	6	AX474735	AX474735 Sequence
c 25	13.2	66.0	96	6	I33379	I33379 Sequence 2
c 26	13	65.0	21	6	AR103525	AR103525 Sequence
27	13	65.0	30	6	E07360	E07360 Primer. 9/1
c 28	12.8	64.0	20	6	AX034348	AX034348 Sequence
c 29	12.8	64.0	21	6	AX096770	AX096770 Sequence
c 30	12.8	64.0	24	6	AR176647	AR176647 Sequence
31	12.8	64.0	31	6	AX248795	AX248795 Sequence
32	12.8	64.0	50	10	MMU41924	U41924 Mus musculu
c 33	12.8	64.0	51	6	AX165632	AX165632 Sequence
c 34	12.8	64.0	82	10	AF152848S1	AF152848 Mus platy
c 35	12.8	64.0	82	10	MSTSPYGI3	AF153222 Mus spret
c 36	12.8	64.0	82	10	MMTSPYGI3	AF153227 Mus maced
c 37	12.8	64.0	84	9	AF267788	AF267788 Homo sapi
38	12.8	64.0	86	5	S76800	S76800 anion excha
c 39	12.8	64.0	99	10	MBTSPYGI3	AF153217 Mus boodu
c 40	12.6	63.0	24	6	AR084744	AR084744 Sequence
c 41	12.6	63.0	30	6	E04878	E04878 Synthetic D
c 42	12.6	63.0	33	6	AX473163	AX473163 Sequence
43	12.6	63.0	51	6	AX159519	AX159519 Sequence
44	12.6	63.0	51	6	AX159520	AX159520 Sequence
45	12.6	63.0	51	6	AX159521	AX159521 Sequence

ALIGNMENTS

RESULT 1
AX350161/c
LOCUS AX350161 23 bp DNA linear PAT 06-FEB-2002
DEFINITION Sequence 20 from Patent WO0202775.
ACCESSION AX350161
VERSION AX350161.1 GI:18615835
KEYWORDS .
SOURCE synthetic construct.
ORGANISM synthetic construct
artificial sequences.
REFERENCE 1
AUTHORS Boehm,T. and Dear,N.T.
TITLE Calpain protease 12
JOURNAL Patent: WO 0202775-A 20 10-JAN-2002;
BASF AKTIENGESELLSCHAFT (DE)

```
FEATURES             Location/Qualifiers
     source          1. .23
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="Capn5-Primer"
BASE COUNT        5 a      7 c      8 g      3 t
ORIGIN

  Query Match            71.0%;  Score 14.2;  DB 6;  Length 23;
  Best Local Similarity   84.2%;  Pred. No. 5.3e+04;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 CCACAGGGCTCACCAGGCG 20
          |||||| ||||  ||||||
Db     23 CCACAGTGCTCTGCAGGCG 5


RESULT 2
I91485/c
LOCUS       I91485                    97 bp    DNA     linear   PAT 01-DEC-1998
DEFINITION  Sequence 19 from patent US 5726017.
ACCESSION   I91485
VERSION     I91485.1  GI:3935955
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 97)
  AUTHORS   Lochrie,M.A. and Gold,L.
  TITLE     High affinity HIV-1 gag nucleic acid ligands
  JOURNAL   Patent: US 5726017-A 19 10-MAR-1998;
FEATURES             Location/Qualifiers
     source          1. .97
                     /organism="unknown"
BASE COUNT       29 a     20 c     35 g     13 t
ORIGIN

  Query Match            71.0%;  Score 14.2;  DB 6;  Length 97;
  Best Local Similarity   84.2%;  Pred. No. 4.1e+04;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 CCACAGGGCTCACCAGGCG 20
          | ||||||||||||||  ||
Db     44 CGACAGGGCTCACCATCCG 26


RESULT 3
MUSTCDXBA
LOCUS       MUSTCDXBA                 78 bp    mRNA    linear   ROD 27-APR-1993
DEFINITION  Mouse T cell receptor rearranged delta-chain gene V-J region, 3'
            end.
ACCESSION   M37589
VERSION     M37589.1  GI:201466
KEYWORDS    J-region; T-cell receptor; V-region; delta-chain; processed gene.
SOURCE      Mouse (strain BALB/c) three week old thymocyte, cDNA to mRNA, clone
            TCD-delta-1.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 78)
  AUTHORS   Okazaki,K. and Sakano,H.
  TITLE     Thymocyte circular DNA excised from T cell receptor alpha-delta
            gene complex
  JOURNAL   EMBO J. 7 (6), 1669-1674 (1988)
  MEDLINE   89005054
   PUBMED   2971535
FEATURES             Location/Qualifiers
     source          1. .78
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
BASE COUNT       16 a     23 c     20 g     19 t
ORIGIN

  Query Match            69.0%;  Score 13.8;  DB 10;  Length 78;
  Best Local Similarity   88.2%;  Pred. No. 6.6e+04;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      3 CACAGGGCTCACCAGGC 19
          ||||| |||| ||||||
Db     35 CACAGTGCTCTCCAGGC 51


RESULT 4
MUSTCDXBB
LOCUS       MUSTCDXBB                 78 bp    mRNA    linear   ROD 27-APR-1993
DEFINITION  Mouse T cell receptor rearranged delta-chain gene V-J region, 3'
            end.
ACCESSION   M37590
VERSION     M37590.1  GI:201467
KEYWORDS    J-region; T-cell receptor; V-region; delta-chain; processed gene.
SOURCE      Mouse (strain BALB/c) three week old thymocyte, cDNA to mRNA, clone
            TCD-delta-2.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 78)
  AUTHORS   Okazaki,K. and Sakano,H.
  TITLE     Thymocyte circular DNA excised from T cell receptor alpha-delta
            gene complex
  JOURNAL   EMBO J. 7 (6), 1669-1674 (1988)
  MEDLINE   89005054
   PUBMED   2971535
FEATURES             Location/Qualifiers
     source          1. .78
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
BASE COUNT       16 a     24 c     20 g     18 t
ORIGIN

  Query Match            69.0%;  Score 13.8;  DB 10;  Length 78;
  Best Local Similarity   88.2%;  Pred. No. 6.6e+04;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      3 CACAGGGCTCACCAGGC 19
          ||||| |||| ||||||
Db     35 CACAGTGCTCTCCAGGC 51


RESULT 5
MUSTCAYA
LOCUS       MUSTCAYA                  78 bp    mRNA    linear   ROD 27-APR-1993
DEFINITION  Mouse T cell receptor rearranged alpha-chain gene, V-J region, 3'
            end.
ACCESSION   M37591
VERSION     M37591.1  GI:201193
KEYWORDS    J-region; T-cell receptor; V-region; alpha-chain; processed gene.
SOURCE      Mouse (strain BALB/c) three week old thymocyte, cDNA to mRNA, clone
            TCD-alpha-3.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 78)
  AUTHORS   Okazaki,K. and Sakano,H.
  TITLE     Thymocyte circular DNA excised from T cell receptor alpha-delta
            gene complex
  JOURNAL   EMBO J. 7 (6), 1669-1674 (1988)
  MEDLINE   89005054
   PUBMED   2971535
FEATURES             Location/Qualifiers
     source          1. .78
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
BASE COUNT       16 a     19 c     18 g     25 t
ORIGIN

  Query Match            69.0%;  Score 13.8;  DB 10;  Length 78;
```

```
  Best Local Similarity   88.2%;  Pred. No. 6.6e+04;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      3 CACAGGGCTCACCAGGC 19
          ||||| |||| ||||||
Db     35 CACAGTGCTCTCCAGGC 51


RESULT 6
MUSTCAYB
LOCUS       MUSTCAYB                  78 bp    mRNA    linear   ROD 27-APR-1993
DEFINITION  Mouse T cell receptor rearranged alpha-chain gene, V-J region, 3'
            end.
ACCESSION   M37592
VERSION     M37592.1  GI:201194
KEYWORDS    J-region; T-cell receptor; V-region; alpha-chain; processed gene.
SOURCE      Mouse (strain BALB/c) three week old thymocyte, cDNA to mRNA, clone
            TCD-alpha-9.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 78)
  AUTHORS   Okazaki,K. and Sakano,H.
  TITLE     Thymocyte circular DNA excised from T cell receptor alpha-delta
            gene complex
  JOURNAL   EMBO J. 7 (6), 1669-1674 (1988)
  MEDLINE   89005054
   PUBMED   2971535
FEATURES             Location/Qualifiers
     source          1. .78
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
BASE COUNT       16 a     20 c     18 g     24 t
ORIGIN

  Query Match             69.0%;  Score 13.8;  DB 10;  Length 78;
  Best Local Similarity   88.2%;  Pred. No. 6.6e+04;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      3 CACAGGGCTCACCAGGC 19
          ||||| |||| ||||||
Db     35 CACAGTGCTCTCCAGGC 51


RESULT 7
HUM13COL15
LOCUS       HUM13COL15                97 bp    DNA     linear   PRI 09-NOV-1994
DEFINITION  Human alpha-1 type XIII collagen (COL13A1) gene, exon 15.
ACCESSION   M68981
VERSION     M68981.1  GI:180730
KEYWORDS    alpha-1 type XIII collagen.
SEGMENT     15 of 39
SOURCE      Homo sapiens (tissue library: genomic cosmid) DNA.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 97)
  AUTHORS   Tikka,L., Elomaa,O., Pihlajaniemi,T. and Tryggvason,K.
  TITLE     Human alpha 1 (XIII) collagen gene. Multiple forms of the gene
            transcripts are generated through complex alternative splicing of
            several short exons
  JOURNAL   J. Biol. Chem. 266 (26), 17713-17719 (1991)
  MEDLINE   91373404
   PUBMED   1894651
FEATURES             Location/Qualifiers
     source          1. .97
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /map="10q22"
                     /tissue_lib="genomic cosmid"
     exon            61. .87
                     /gene="COL13A1"
                     /note="G00-119-789"
                     /number=15
                     /evidence=experimental
BASE COUNT       14 a     28 c     18 g     37 t
ORIGIN      chromosome 10.

  Query Match             69.0%;  Score 13.8;  DB 9;  Length 97;
  Best Local Similarity   88.2%;  Pred. No. 6.4e+04;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      2 CCACAGGGCTCACCAGG 18
          |||||||||| ||| ||
Db     55 CCACAGGGCTTACCTGG 71


RESULT 8
AX204305/c
LOCUS       AX204305                  51 bp    DNA     linear   PAT 30-AUG-2001
DEFINITION  Sequence 411 from Patent WO0148245.
ACCESSION   AX204305
VERSION     AX204305.1  GI:15393825
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 51)
  AUTHORS   Shimkets,R.A. and Leach,M.
  TITLE     Nucleic acids containing single nucleotide polymorphisms and
            methods of use thereof
  JOURNAL   Patent: WO 0148245-A 411 05-JUL-2001;
            Curagen Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .51
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     variation       26
                     /note="single nucleotide polymorphism
                     Accession number cg25236776"
BASE COUNT        5 a     26 c     13 g      7 t
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 6;  Length 51;
  Best Local Similarity   80.0%;  Pred. No. 8.9e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 ACCACAGGGCTCACCAGGCG 20
          | ||| ||||||| |||||
Db     29 AGCACGGGGCTCAGTAGGCG 10


RESULT 9
REO2L3A/c
LOCUS       REO2L3A                   70 bp ds-RNA     linear   VRL 03-AUG-1993
DEFINITION  Reovirus serotype 2 L3 gene, major core protein lambda-1, 5' end.
ACCESSION   J02305
VERSION     J02305.1  GI:333655
KEYWORDS    major core protein; major core protein lambda-1.
SOURCE      Reovirus (D5/Jones strain; serotype 2) RNA.
  ORGANISM  Mammalian orthoreovirus 2
            Viruses; dsRNA viruses; Reoviridae; Orthoreovirus; Mammalian
            orthoreoviruses.
REFERENCE   1  (bases 1 to 70)
  AUTHORS   Gaillard,R.K., Li,J.K., Keene,J.D. and Joklik,W.K.
  TITLE     The sequence at the termini of four genes of the three reovirus
            serotypes
  JOURNAL   Virology 121 (2), 320-326 (1982)
  MEDLINE   83017877
   PUBMED   7123853
COMMENT     [1] compares given sequence with 5' ends of serotype 1 and 3 L3
            genes. [1] sequence was deduced. Plus strand is shown.
FEATURES             Location/Qualifiers
     source          1. .70
                     /organism="Mammalian orthoreovirus 2"
```

```
                     /db_xref="taxon:10885"
     CDS             14. .>70
                     /note="major core protein lambda-1"
                     /codon_start=1
                     /protein_id="AAA47244.1"
                     /db_xref="GI:333656"
                     /translation="MKRIPRKTRGKSSGKGNDS"
BASE COUNT       23 a     12 c     21 g     14 t
ORIGIN      Unreported.

  Query Match             68.0%;  Score 13.6;  DB 14;  Length 70;
  Best Local Similarity   80.0%;  Pred. No. 8.4e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 ACCACAGGGCTCACCAGGCG 20
          ||||| || || |||| |||
Db     55 ACCACTGGACTTACCACGCG 36


RESULT 10
AX145934
LOCUS       AX145934                 21 bp    DNA     linear   PAT 31-MAY-2001
DEFINITION  Sequence 125 from Patent WO0134840.
ACCESSION   AX145934
VERSION     AX145934.1  GI:14284452
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 21)
  AUTHORS   Au,K.G., Chen,J.G., Patil,N. and Thomas,D.
  TITLE     Genetic compositions and methods
  JOURNAL   Patent: WO 0134840-A 125 17-MAY-2001;
            GLAXO GROUP LIMITED (GB) ; Affymetrix, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .21
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     variation       1. .21
                     /note="'n' represents a polymorphic base"
BASE COUNT        5 a      7 c      7 g      1 t      1 others
ORIGIN

  Query Match             67.0%;  Score 13.4;  DB 6;  Length 21;
  Best Local Similarity   87.5%;  Pred. No. 1.3e+05;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      3 CACAGGGCTCACCAGG 18
          |||||||| | |||||
Db      1 CACAGGGCACNCCAGG 16


RESULT 11
A48205
LOCUS       A48205                   20 bp    DNA     linear   PAT 07-MAR-1997
DEFINITION  Sequence 4 from Patent EP0698665.
ACCESSION   A48205
VERSION     A48205.1  GI:2302052
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 20)
  AUTHORS   Heinz,F.X., Kunz,C.P. and Mandl,C.D.
  TITLE     Infectious Tick-Borne Encephalitis virus cDNA clone, recombinant
            vaccine derived therefrom and preparation thereof
  JOURNAL   Patent: EP 0698665-A 4 28-FEB-1996;
            IMMUNO AG (AT)
COMMENT     Other publication HU 72840 960528
            Other publication CZ 9501850 960214
            Other publication FI 953581 960128
            Other publication DE 4426622 960222.
FEATURES             Location/Qualifiers
     source          1. .20
                     /organism="unidentified"
                     /db_xref="taxon:32644"
BASE COUNT        6 a      9 c      4 g      1 t
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 6;  Length 20;
  Best Local Similarity   83.3%;  Pred. No. 1.6e+05;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 ACCACAGGGCTCACCAGG 18
          ||||||  || |||||||
Db      3 ACCACATAGCGCACCAGG 20


RESULT 12
A57298
LOCUS       A57298                   20 bp    DNA     linear   PAT 03-MAR-1998
DEFINITION  Sequence 4 from Patent WO9630521.
ACCESSION   A57298
VERSION     A57298.1  GI:3713193
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 20)
  AUTHORS   Heinz,F.X., Kunz,C. and Mandl,C.
  TITLE     INFECTIOUS DNA CLONE OF THE TICK-BORNE ENCEPHALITIS (TBE) VIRUS;
            RECOMBINANT VACCINE DERIVED FROM THE SAID CLONE AND PRODUCTION OF
            THE SAID VACCINE; AND A PHARMACEUTICAL PRODUCT CONTAINING A
            REPLICABLE NUCLEIC ACID
  JOURNAL   Patent: WO 9630521-A 4 03-OCT-1996;
            IMMUNO AG (AT)
COMMENT     Other publication AU 5398596 961016
            Other publication DE 19512142 961002.
FEATURES             Location/Qualifiers
     source          1. .20
                     /organism="unidentified"
                     /db_xref="taxon:32644"
BASE COUNT        6 a      9 c      4 g      1 t
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 6;  Length 20;
  Best Local Similarity   83.3%;  Pred. No. 1.6e+05;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 ACCACAGGGCTCACCAGG 18
          ||||||  || |||||||
Db      3 ACCACATAGCGCACCAGG 20


RESULT 13
AR076724
LOCUS       AR076724                 20 bp    DNA     linear   PAT 30-AUG-2000
DEFINITION  Sequence 89 from patent US 5959096.
ACCESSION   AR076724
VERSION     AR076724.1  GI:10003470
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 20)
  AUTHORS   Bennett,C.Frank. and Dean,N.
  TITLE     Antisense oligonucleotides against human protein kinase C
  JOURNAL   Patent: US 5959096-A 89 28-SEP-1999;
FEATURES             Location/Qualifiers
     source          1. .20
                     /organism="unknown"
BASE COUNT        4 a      11 c     4 g      1 t
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 6;  Length 20;
```

```
  Best Local Similarity   83.3%;  Pred. No. 1.6e+05;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 CCACAGGGCTCACCAGGC 19
          || |||||| |||||| |
Db      1 CCCCAGGGCCCACCAGTC 18


RESULT 14
AR076738
LOCUS       AR076738                  20 bp    DNA     linear   PAT 30-AUG-2000
DEFINITION  Sequence 103 from patent US 5959096.
ACCESSION   AR076738
VERSION     AR076738.1  GI:10003484
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 20)
  AUTHORS   Bennett,C.Frank. and Dean,N.
  TITLE     Antisense oligonucleotides against human protein kinase C
  JOURNAL   Patent: US 5959096-A 103 28-SEP-1999;
FEATURES             Location/Qualifiers
     source          1. .20
                     /organism="unknown"
BASE COUNT        3 a      12 c      4 g      1 t
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 6;  Length 20;
  Best Local Similarity   83.3%;  Pred. No. 1.6e+05;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 CCACAGGGCTCACCAGGC 19
          || |||||| |||||| |
Db      2 CCCCAGGGCCCACCAGTC 19


RESULT 15
AR182781
LOCUS       AR182781                  20 bp    DNA     linear   PAT 20-APR-2002
DEFINITION  Sequence 89 from patent US 6339066.
ACCESSION   AR182781
VERSION     AR182781.1  GI:20225988
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 20)
  AUTHORS   Bennett,C.Frank., Dean,N.M., Cook,P.Dan. and Hoke,G.
  TITLE     Antisense oligonucleotides which have phosphorothioate linkages of
            high chiral purity and which modulate .beta.I, .beta.II, .gamma.,
            .delta., .EPSILON., .zeta. and .eta. isoforms of human protein
            kinase C
  JOURNAL   Patent: US 6339066-A 89 15-JAN-2002;
FEATURES             Location/Qualifiers
     source          1. .20
                     /organism="unknown"
BASE COUNT        4 a      11 c      4 g      1 t
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 6;  Length 20;
  Best Local Similarity   83.3%;  Pred. No. 1.6e+05;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 CCACAGGGCTCACCAGGC 19
          || |||||| |||||| |
Db      1 CCCCAGGGCCCACCAGTC 18


Search completed: December  3, 2002, 22:23:58
Job time : 358.3 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 06:29:45 ; Search time 94.1 Seconds
(without alignments)
478.639 Million cell updates/sec

Title: US-09-296-264-30
Perfect score: 20
Sequence: 1 accacagggctcaccaggcg 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 2390332

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	20	100.0	20	21	AAZ31460	Human neuropilin m
c	2	14.2	71.0	23	24	ABA99783	Murine capn5 Set 1
c	3	14.2	71.0	93	22	AAH84149	Human cell death p
c	4	13.8	69.0	24	21	AAA37979	PCR primer Q used
	5	13.8	69.0	65	24	ABN29553	Rat spliced transc
	6	13.6	68.0	48	21	AAA46995	Oligonucleotide us
c	7	13.6	68.0	51	22	AAH79796	Human DNA containi
	8	13.4	67.0	50	22	AAL34460	Human SNP oligonuc
	9	13.2	66.0	20	16	AAQ97960	PNA oligomer targe
	10	13.2	66.0	20	16	AAQ84237	PKC-epsilon coding
	11	13.2	66.0	20	17	AAT44474	Tick-borne encepha
	12	13.2	66.0	20	17	AAT15828	TBE virus strain N
	13	13.2	66.0	20	20	AAZ27354	Human protein kina
	14	13.2	66.0	20	20	AAZ27368	Human protein kina
	15	13.2	66.0	20	20	AAX78612	Human PKC-epsilon
	16	13.2	66.0	20	20	AAX78626	Human PKC-epsilon
	17	13.2	66.0	20	20	AAX83747	Human protein kina
	18	13.2	66.0	20	20	AAX83704	Human protein kina
	19	13.2	66.0	20	20	AAX22664	Human protein kina
	20	13.2	66.0	20	20	AAX22650	Human protein kina
	21	13.2	66.0	20	20	AAX19229	Human PKC-epsilon
	22	13.2	66.0	20	20	AAX19215	Human PKC-epsilon
	23	13.2	66.0	20	24	ABL90942	Human protein kina
	24	13.2	66.0	20	24	ABL90956	Human protein kina
c	25	13.2	66.0	21	22	AAH45861	Porcine Fas ligand
c	26	13.2	66.0	21	22	AAF97183	Human gene single
c	27	13.2	66.0	50	13	AAQ31970	P53 binding site o
	28	13.2	66.0	50	16	AAQ94298	Reverse transcript
	29	13.2	66.0	50	16	AAQ95038	Human hippocampal
	30	13.2	66.0	60	24	ABN40770	Human spliced tran
c	31	13.2	66.0	60	24	ABN42122	Human spliced tran
c	32	13.2	66.0	65	24	ABN28221	Rat spliced transc
c	33	13.2	66.0	75	24	ABL51903	cGMP-hammerhead RN
	34	13.2	66.0	90	22	ABA35426	Probe #13892 for g
c	35	13	65.0	21	20	AAZ18399	Polymorphic fragme
c	36	13	65.0	21	21	AAA80301	Human ASTH1J 5' re
	37	13	65.0	30	15	AAQ76206	Primer to amplify
c	38	12.8	64.0	20	18	AAT93926	Primer for exon 19
c	39	12.8	64.0	20	21	AAA50417	Collagen I gene an
	40	12.8	64.0	21	23	AAH88921	Human polymorphic
	41	12.8	64.0	24	21	AAZ60718	PCR primer used to
	42	12.8	64.0	29	24	ABL60038	X. albilineans xab
c	43	12.8	64.0	33	22	AAC89168	p-HYDE sequence fr
c	44	12.8	64.0	43	24	ABN84312	Rhinovirus specifi
c	45	12.8	64.0	51	23	ABL00836	Human amino acid c

ALIGNMENTS

RESULT 1
AAZ31460
ID AAZ31460 standard; DNA; 20 BP.
XX
AC AAZ31460;
XX
DT 07-FEB-2000 (first entry)
XX
DE Human neuropilin mRNA specific antisense oligo GTI3631.
XX
KW Neuropilin; human; growth; metastasis; tumor; neovascularisation;
KW cancer; papilloma; diabetic retinopathy; antisense; ss.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9955855-A2.
XX
PD 04-NOV-1999.
XX
PF 23-APR-1999; 99WO-CA00324.
XX
PR 23-APR-1998; 98US-0082791.
XX
PA (GENE-) GENESENSE TECHNOLOGIES INC.
XX
PI Wright JA, Young AH, Lee YS;
XX
DR WPI; 2000-023357/02.
XX
PT Antisense oligonucleotides that inhibit neuropilin expression, useful
PT for treating cancer -

XX
PS Claim 4; Page 17; 57pp; English.
XX
CC Sequences AAZ31431-460 represent antisense oligonucleotides which
CC inhibit human neuropilin expression. The antisense oligonucleotides can
CC be used to inhibit the growth or metastasis of a mammalian tumor and
CC inhibit neovascularisation. The oligonucleotides may be used to treat
CC various forms of cancers or tumors, such as sarcomas, melanomas,
CC adenomas, carcinomas of solid tissue, hypoxic tumors, squamous cell
CC carcinomas of the mouth, throat, larynx and lung, genitourinary cancers
CC such as cervical and bladder cancer, hematopoietic cancers, colon cancer,
CC breast cancer, pancreatic cancer, renal cancer, brain cancer, skin
CC cancer, liver cancer, head and neck cancers, and nervous system cancers,
CC as well as benign lesions such as papillomas. The methods may be used to
CC treat neovascularisation disorders such as diabetic retinopathy, and
CC retinopathy of prematurity and age related macular degeneration.
XX
SQ Sequence 20 BP; 5 A; 8 C; 6 G; 1 T; 0 other;

Query Match 100.0%; Score 20; DB 21; Length 20;
Best Local Similarity 100.0%; Pred. No. 4.8;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 ACCACAGGGCTCACCAGGCG 20
          ||||||||||||||||||||
Db      1 ACCACAGGGCTCACCAGGCG 20
```

RESULT 2
ABA99783/c
ID ABA99783 standard; DNA; 23 BP.
XX
AC ABA99783;
XX
DT 11-JUN-2002 (first entry)
XX
DE Murine capn5 Set 1 PCR primer SEQ ID NO 20.
XX
KW Calpain protease; murine; gene therapy; PCR; primer; screening;
KW diagnosis; capn12; capn5; ss.
XX
OS Mus sp.
XX
PN DE10031932-A1.
XX
PD 10-JAN-2002.
XX
PF 30-JUN-2000; 2000DE-1031932.
XX
PR 30-JUN-2000; 2000DE-1031932.
XX
PA (BADI) BASF AG.
XX
PI Not given;
XX
DR WPI; 2002-115441/16.
XX
PT New calpain protein 12 with cysteine protease activity, useful for
PT treating specific deficiency disorders -
XX
PS Example 7; Page 8; 36pp; German.
XX
CC This invention describes a novel murine calpain protease 12 (capn12).
CC The calpain protease of the invention, related proteins and nucleic acid
CC that encodes it, are useful for treatment (including gene therapy) of
CC diseases associated with insufficient expression of the calpain protease.
CC The protein is also used to screen for calpain protein effectors and to
CC raise specific immunoglobulins (Ig) useful for diagnosis. Also the
CC polynucleotide encoding capn12 is useful, e.g. as primers and probes, for
CC diagnosis of diseases, or predisposition to them, and for recombinant
CC production of capn12. This sequence represents a PCR primer used in
CC the amplification of the murine calpain protease, capn5 described in the
CC disclosure of the invention.
XX
SQ Sequence 23 BP; 5 A; 7 C; 8 G; 3 T; 0 other;

Query Match 71.0%; Score 14.2; DB 24; Length 23;
Best Local Similarity 84.2%; Pred. No. 2.3e+03;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy      2 CCACAGGGCTCACCAGGCG 20
          |||||| ||||  ||||||
Db     23 CCACAGTGCTCTGCAGGCG 5
```

RESULT 3
AAH84149/c
ID AAH84149 standard; cDNA; 93 BP.
XX
AC AAH84149;
XX
DT 21-SEP-2001 (first entry)
XX
DE Human cell death protective cDNA clone CNI-00712 ORF4, SEQ:33.
XX
KW Cell death protective; apoptosis; necrosis; human; drug screening;
KW cell death-associated disorder; central nervous system disorder;
KW psychiatric disorder; neurological disorder; ischaemia-related disorder;
KW stroke; cerebral infarction; ischaemic encephalopathy;
KW neurodegenerative disorder; Alzheimer's disease; Huntington's disease;
KW Parkinson's disease; infection; meningitis; malaria; trypanosomiasis;
KW vascular disease; opthalmological disorder; diabetic retinopathy;
KW macular degeneration; hypertension; myocardial infarction;
KW atherosclerosis; respiratory disorder; asthma; transgenic animal;
KW chronic obstructive pulmonary disease; neoplastic condition; cancer;
KW benign tumour; anaemia; gastrointestinal disorder; gastritis;
KW ulcerative colitis; liver disease; biliary cirrhosis; kidney disorder;
KW glomerulonephritis; cystitis; endometriosis; endocrine disorder;
KW Grave's disease; Hashimoto's thyroiditis; skin condition; dermatitis;
KW urticaria; immune disorder; acquired immunodeficiency syndrome; AIDS;
KW open reading frame; ORF; ss.
XX
OS Homo sapiens.
XX
PN WO200145638-A2.
XX
PD 28-JUN-2001.
XX
PF 11-DEC-2000; 2000WO-US33547.
XX
PR 14-DEC-1999; 99US-0461697.
XX
PA (COGE-) COGENT NEUROSCIENCE INC.
XX
PI Lo DC, Barney S, Thomas MB, Portbury SD, Puranam K, Katz LC;
XX
DR WPI; 2001-390297/41.
DR P-PSDB; AAG98623.
XX
PT Novel protective sequence polynucleotides and polypeptides, used to
PT identify modulators of their expression and activity, which are used in
PT to treat central nervous system conditions, diseases and disorders -
XX
PS Claim 2; Fig 5D; 325pp; English.
XX
CC Sequences AAH84132-AAH84370 represent human nucleic acid sequences which
CC protect against cell death (i.e., apoptosis or necrosis). Sequences
CC AAH84132, AAH84145, AAH84170, AAH84201, AAH84210, AAH84226, AAH84265,
CC AAH84281, AAH84315 and AAH84367 represent 10 full-length cDNA clones,
CC while the remaining nucleic acid sequences within the range given above
CC represent the open reading frames (ORFs) of these cDNA clones. Sequences
CC AAG98610-AAG98829 represent the polypeptides encoded by the cell death
CC protective ORFs. The cell death protective cDNA clones are able to
CC prevent, delay or reverse progression through the apoptotic or necrotic
CC pathways when injected into a cell predisposed to or undergoing cell
CC death. The cell death protective nucleic acids and polypeptides can be

```
CC   used in the diagnosis and treatment of disorders associated with cell
CC   death, and to screen for compounds which modulate their activity or
CC   expression. Such modulators, preferably a small organic molecule, an
CC   antibody, a ribozyme, or an antisense molecule, can also be used to treat
CC   cell death-related diseases. Such diseases include those associated with
CC   the central nervous system including psychiatric or neurological
CC   disorders, especially ischaemia-related conditions such as strokes, and
CC   also includes neurodegenerative disorders such as Alzheimer's disease,
CC   Huntington's disease, or Parkinson's disease. The modulators may also be
CC   used to treat infections such as meningitis, malaria, or trypanosomiasis;
CC   vascular diseases such as ischaemic encephalopathy or cerebral
CC   infarction; eye conditions such as diabetic retinopathy or macular
CC   degeneration; hypertension; myocardial infarction; atherosclerosis;
CC   respiratory conditions such as asthma or chronic obstructive pulmonary
CC   disease; neoplastic conditions such as cancers or benign tumours; blood
CC   cell conditions such as anaemia; gastrointestinal conditions such as
CC   gastritis or ulcerative colitis; liver conditions such as biliary
CC   cirrhosis; kidney disorders such as glomerulonephritis; cystitis;
CC   endometriosis; endocrine disorders such as Grave's disease or Hashimoto's
CC   thyroiditis; skin conditions such as dermatitis or urticaria; or immune
CC   system disorders such as acquired immunodeficiency syndrome (AIDS). The
CC   nucleic acids may additionally be used to generate animal models of
CC   cell death-associated disorders. The present sequence represents a
CC   cell death protective ORF.
XX
SQ   Sequence 93 BP; 18 A; 34 C; 27 G; 14 T; 0 other;

  Query Match             71.0%;  Score 14.2;  DB 22;  Length 93;
  Best Local Similarity   84.2%;  Pred. No. 2.4e+03;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 CCACAGGGCTCACCAGGCG 20
           |||| |||||| |||| ||
Db      28 CCACTGGGCTCCCCAGTCG 10


RESULT 4
AAA37979/c
ID   AAA37979 standard; DNA; 24 BP.
XX
AC   AAA37979;
XX
DT   18-AUG-2000  (first entry)
XX
DE   PCR primer Q used in carrot promoter containing vector construction.
XX
KW   Carrot; promoter; terminator; transgenic plant; breeding; fertility;
KW   PCR primer; ss.
XX
OS   Daucus carota.
XX
PN   WO200020613-A1.
XX
PD   13-APR-2000.
XX
PF   28-SEP-1999;   99WO-JP05303.
XX
PR   02-OCT-1998;   98JP-0281124.
XX
PA   (SUMO ) SUMITOMO CHEM CO LTD.
XX
PI   Nishikawa S,  Oeda K;
XX
DR   WPI; 2000-303791/26.
XX
PT   New Plant promoters and terminators from Daucus carota L., useful in
PT   plant breeding, for e.g. controlling fertilities of plants -
XX
PS   Example 6; Page 43; 81pp; English.
XX
CC   This sequence represents a PCR primer used in the construction of a
CC   vector containing the promoter of the invention. The invention relates to
CC   plant promoters and terminators from Daucus carota L. which are capable
CC   of expressing a gene of interest in plants. The invention also includes a
CC   chimeric gene characterized in that it comprises the promoter and a
CC   desired gene linked to each other in the form capable of functioning. A
CC   method of producing a transformant comprises introducing the promoter,
CC   the chimeric gene or a vector comprising the promoter and a desired
CC   gene or terminator sequence into a host cell. The plant promoters and
CC   terminators are useful in plant breeding, for e.g. fertilities of plants
CC   may be controlled by expressing, in the host cells, a sense or antisense
CC   gene of a male sterility related gene such as S-locus-specific RNase
CC   gene.
XX
SQ   Sequence 24 BP; 4 A; 5 C; 9 G; 6 T; 0 other;

  Query Match             69.0%;  Score 13.8;  DB 21;  Length 24;
  Best Local Similarity   88.2%;  Pred. No. 3.5e+03;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 ACCACAGGGCTCACCAG 17
           ||||||||| || ||||
Db      22 ACCACAGGGTTCCCCAG 6


RESULT 5
ABN29553
ID   ABN29553 standard; DNA; 65 BP.
XX
AC   ABN29553;
XX
DT   15-JUL-2002  (first entry)
XX
DE   Rat spliced transcript detection oligonucleotide SEQ ID NO:2301.
XX
KW   Human; mouse; rat; splice transcript; detection; RNA transcript;
KW   splice variant; transcriptome; oligonucleotide library; ss.
XX
OS   Rattus norvegicus.
XX
PN   WO200210449-A2.
XX
PD   07-FEB-2002.
XX
PF   20-JUL-2001; 2001WO-IB01903.
XX
PR   28-JUL-2000; 2000US-221607P.
PR   02-MAY-2001; 2001US-287724P.
XX
PA   (COMP-) COMPUGEN INC.
XX
PI   Shoshan A,  Wasserman A,  Mintz E,  Mintz L,  Faigler S;
XX
DR   WPI; 2002-257383/30.
XX
PT   New oligonucleotide libraries comprising oligonucleotides which
PT   selectively hybridize to mRNAs transcribed from a transcription unit of
PT   a genome, useful for detecting tissue-, pathology-, and
PT   developmental-specific genes  -
XX
PS   Example 1; SEQ ID 2301; 47pp; English.
XX
CC   The present invention describes oligonucleotide libraries for detecting
CC   messenger RNAs that populate a (sub-)transcriptome, where the
CC   (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC   transcription units that populate a genome. The library comprises
CC   several oligonucleotides, each capable of hybridising selectively to a
CC   set of messenger RNAs transcribed from a given transcription unit of
CC   the genome, which encodes one or more messenger RNA splice variants.
CC   The oligonucleotide libraries are useful for detecting mRNAs from a
CC   biological sample, in expression profiling studies, in qualitatively or
CC   quantitatively characterising the corresponding transcriptome, and in
CC   detecting RNA transcripts and splice variants of human or animal
CC   transcriptomes. The libraries may also be used as specialised mini
CC   libraries to detect transcripts of a sub-transcriptome under a
CC   particular biological or pathological state, and so allowing the
```

```
CC   detection of tissue- and pathology-specific genes such as those genes
CC   only expressed in specific tissue under a specific pathological
CC   condition; to detect developmental specific genes; and to detect RNA
CC   transcripts and splice variants of a transcriptome of a patient suffering
CC   from a particular disorder. ABN27253 to ABN59589 represent
CC   oligonucleotide sequences from rats, humans and mice, which are used in
CC   the exemplification of the present invention.
CC   N.B. The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 65 BP; 15 A; 14 C; 19 G; 17 T; 0 other;

  Query Match             69.0%;  Score 13.8;  DB 24;  Length 65;
  Best Local Similarity   88.2%;  Pred. No. 3.6e+03;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 ACCACAGGGCTCACCAG 17
           ||||||| ||||| |||
Db      15 ACCACAGTGCTCAACAG 31


RESULT 6
AAA46995
ID   AAA46995 standard; cDNA; 48 BP.
XX
AC   AAA46995;
XX
DT   03-OCT-2000  (first entry)
XX
DE   Oligonucleotide used for in situ analysis of PRO1017 DNA.
XX
KW   PRO201; PRO292; PRO327; PRO1265;  PRO344; PRO343; PRO347; PRO357;
KW   PRO715; PRO1017; PRO1112; PRO509; PRO853; PRO882; tumour cell; probe;
KW   tumourigenesis; cancer; neoplastic cell growth; cell proliferation; ss.
XX
OS   Homo sapiens.
XX
PN   WO200037640-A2.
XX
PD   29-JUN-2000.
XX
PF   16-DEC-1999;   99WO-US30095.
XX
PR   22-DEC-1998;   98US-0113296.
PR   08-MAR-1999;   99WO-US05028.
PR   02-JUN-1999;   99WO-US12252.
PR   01-SEP-1999;   99WO-US20111.
PR   15-SEP-1999;   99WO-US21090.
PR   30-NOV-1999;   99WO-US28313.
PR   30-NOV-1999;   99WO-US28409.
PR   01-DEC-1999;   99WO-US28301.
PR   02-DEC-1999;   99WO-US28565.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Botstein D,  Goddard A,  Gurney AL,  Hillan K,  Lawrence DA,  Roy MA;
PI   Wood WI;
XX
DR   WPI; 2000-452188/39.
XX
PT   New anti-polypeptide antibody useful in the treatment and diagnosis of
PT   neoplastic cell growth and proliferation -
XX
PS   Example 18; Page 162; 220pp; English.
XX
CC   The present sequence was used for in situ analysis of DNA encoding
CC   a novel human polypeptide. The specification describes novel polypeptides
CC   designated PRO201, PRO292, PRO327, PRO1265, PRO344, PRO343, PRO347,
CC   PRO357, PRO715, PRO1017, PRO1112, PRO509, PRO853 and PRO882. These
CC   genes are amplified in the genome of tumour cells. The polypeptides
CC   are believed to contribute to tumourigenesis. The polypeptides are
CC   useful target for the identification of certain cancers, and may act
CC   as predictors of the prognosis of tumour treatment. Antibodies against
CC   these polypeptides are useful in the treatment and diagnosis of
CC   neoplastic cell growth and proliferation in mammals.
XX
SQ   Sequence 48 BP; 13 A; 10 C; 18 G; 7 T; 0 other;

  Query Match             68.0%;  Score 13.6;  DB 21;  Length 48;
  Best Local Similarity   80.0%;  Pred. No. 4.4e+03;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 ACCACAGGGCTCACCAGGCG 20
           || | ||||| || ||||||
Db      18 ACTATAGGGCGCAGCAGGCG 37


RESULT 7
AAH79796/c
ID   AAH79796 standard; DNA; 51 BP.
XX
AC   AAH79796;
XX
DT   19-SEP-2001  (first entry)
XX
DE   Human DNA containing single nucleotide polymorphism SEQ ID NO. 411.
XX
KW   Human; single nucleotide polymorphism; SNP; angiopoietin;
KW   4-hydroxybutyrate; dehydrogenase; protein therapy;
KW   adenosine triphosphate-dependent RNA helicase;
KW   major histocompatibility complex Class I histocompatibility antigen; MHC;
KW   phosphoglycerate kinase; immunosuppressive; immunostimulatory;
KW   antirheumatic; antisclerotic; antidiabetic; antiinflammatory; cytostatic;
KW   antileukemic; neuroprotective; antimicrobial; gene therapy; vaccine; ds.
XX
OS   Homo sapiens.
XX
PN   WO200148245-A2.
XX
PD   05-JUL-2001.
XX
PF   27-DEC-2000; 2000WO-US35346.
XX
PR   27-DEC-1999;   99US-0472688.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Shimkets RA,  Leach M;
XX
DR   WPI; 2001-418297/44.
XX
PT   Polymorphic nucleic acids encoding e.g. angiopoietin, dehydrogenase,
PT   adenosine triphosphate-dependent RNA helicase and/or phosphoglycerate
PT   kinase, useful for diagnosing and treating, e.g. cancer, autoimmune
PT   diseases and infections -
XX
PS   Claim 1; Page 175; 484pp; English.
XX
CC   The invention relates to nucleic acids (AAH79386-AAH80036) encoding
CC   polymorphic variants of proteins (AAG98010-AAG98238) related to
CC   angiopoietin, 4-hydroxybutyrate, dehydrogenase, adenosine triphosphate
CC   (ATP)-dependent RNA helicase, major histocompatibility complex (MHC)
CC   Class I histocompatibility antigen and/or phosphoglycerate kinase. These
CC   nucleic acid single nucleotide polymorphisms (SNPs) and the encoded
CC   proteins have potential immunosuppressive, immunostimulatory,
CC   antirheumatic, antisclerotic, antidiabetic, antiinflammatory, cytostatic,
CC   antileukemic, neuroprotective and antimicrobial activity and may be
CC   useful in gene/protein therapy, vaccines, modulation of the expression
CC   and activity of proteins related to angiopoietin, 4-hydroxybutyrate,
CC   dehydrogenase, adenosine triphosphate (ATP)-dependent RNA helicase,
CC   major histocompatibility complex (MHC) Class I histocompatibility antigen
CC   and/or phosphoglycerate kinase. Disorders that may be prevented,
CC   diagnosed and/or treated by the above methods include multifactorial
CC   diseases with a genetic component, such as autoimmune diseases (e.g.
CC   rheumatoid arthritis, multiple sclerosis, diabetes, systemic lupus
```

```
CC   erytheromatosus and Grave's disease), inflammation, cancer (e.g. cancers
CC   of the bladder, brain, breast, colon and kidney, leukemia), diseases of
CC   the nervous system, an infection of pathogenic organisms. They may also
CC   be used to alter phenotypic traits such as longevity, appearance,
CC   strength, speed and endurance.
XX
SQ   Sequence 51 BP; 5 A; 26 C; 13 G; 7 T; 0 other;

  Query Match             68.0%;  Score 13.6;  DB 22;  Length 51;
  Best Local Similarity   80.0%;  Pred. No. 4.4e+03;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       1 ACCACAGGGCTCACCAGGCG 20
           | ||| |||||||  |||||
Db      29 AGCACGGGGCTCAGTAGGCG 10


RESULT 8
AAL34460
ID   AAL34460 standard; DNA; 50 BP.
XX
AC   AAL34460;
XX
DT   24-JAN-2002  (first entry)
XX
DE   Human SNP oligonucleotide #7668.
XX
KW   Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
KW   neuroprotective; antimicrobial; gene therapy; vaccine; amylase; cancer;
KW   amyloid protein; angiopoietin; apoptosis related protein; cadherin;
KW   cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
KW   complement related protein; cytochrome; kinesin; cytokine; interferon;
KW   interleukin; G-protein coupled receptor; thioesterase; inflammation;
KW   multifactorial disease; autoimmune disease; infection;
KW   nervous system disease; ss.
XX
OS   Homo sapiens.
XX
PN   WO200147944-A2.
XX
PD   05-JUL-2001.
XX
PF   28-DEC-2000; 2000WO-US35498.
XX
PR   28-DEC-1999;   99US-0173419.
PR   27-DEC-2000; 2000US-0173419.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Shimkets RA,  Leach M;
XX
DR   WPI; 2001-465210/50.
XX
PT   Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
PT   oncogenes and histones, useful for diagnosing and treating, e.g.
PT   cancer, autoimmune diseases and infections -
XX
PS   Claim 1; Page 3603; 4143pp; English.
XX
CC   The present invention relates to oligonucleotides encoding polymorphic
CC   variants of proteins related to amylases, amyloid proteins, angiopoietin,
CC   apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,
CC   histones, kinases, colony stimulating factors, complement related
CC   proteins, cytochromes, kinesins, cytokines, interferons, interleukins,
CC   G-protein coupled receptors and thioesterases. The present sequence is
CC   one such oligonucleotide. The oligonucleotides and the peptides encoded
CC   by them may be used in the prevention, diagnosis and treatment of
CC   diseases associated with inappropriate expression of the proteins listed
CC   above. Disorders that may be prevented, diagnosed and/or treated include
CC   multifactorial diseases with a genetic component, such as autoimmune
CC   diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,
CC   systemic lupus erythromatosus and Grave's disease), inflammation, cancer
CC   (e.g. cancers of the bladder, brain, breast, colon and kidney,
CC   leukaemia), diseases of the nervous system and an infection of pathogenic
CC   organisms.
XX
SQ   Sequence 50 BP; 14 A; 12 C; 12 G; 12 T; 0 other;

  Query Match             67.0%;  Score 13.4;  DB 22;  Length 50;
  Best Local Similarity   93.3%;  Pred. No. 5.4e+03;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 ACCACAGGGCTCACC 15
           ||||||||| |||||
Db      36 ACCACAGGGATCACC 50


RESULT 9
AAQ97960
ID   AAQ97960 standard; DNA; 20 BP.
XX
AC   AAQ97960;
XX
DT   18-OCT-1995  (first entry)
XX
DE   PNA oligomer targetting coding region of PKC-epsilon.
XX
KW   Peptide nucleic acid; PNA; PKC-alpha; protein kinase C; ss;
KW   cell proliferation; cell differentiation; isozyme; antisense;
KW   triple helix; cancer; psoriasis; inflammation.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   misc_feature    1..20
FT                   /*tag= a
FT                   /note= "at least one (and preferably all) of
FT                          the backbone subunits are composed of N-acetyl
FT                          N-(2-aminoethyl)glycine peptide residues, the
FT                          nucleobase being attached covalently to the
FT                          acetyl group and the peptide linkage being
FT                          formed by condensation of the glycine
FT                          carboxy group of one residue with the amino
FT                          group of the 2-aminoethyl moiety in the next
FT                          residue"
XX
PN   WO9503833-A.
XX
PD   09-FEB-1995.
XX
PF   28-JUL-1994;   94WO-US08465.
XX
PR   29-JUL-1993;   93US-0099098.
XX
PA   (ISIS-) ISIS PHARM INC.
XX
PI   Dean NM;
XX
DR   WPI; 1995-082040/11.
XX
PT   New peptide nucleic acid oligomers specific for protein kinase C
PT   isozyme(s) - useful as anti:sense molecules for treating PKC
PT   mediated disease, e.g. cancer, psoriasis and inflammation
XX
PS   Claim 38; Page 274; 287pp; English.
XX
CC   New peptide nucleic acid (PNA) oligomers are provided which (a) consist
CC   of naturally occurring nucleobases covalently bound to a polyamide
CC   backbone and (b) hybridise to the translation initiation AUG region,
CC   coding region, 5' untranslated region (5' UTR) or 3' untranslated region
CC   (3' UTR) of PKC-alpha or its isoforms. The PNAs can be used to target
CC   RNA and single stranded DNA (ssDNA) to produce antisense-type gene
CC   regulation moieties. They inhibit expression of PKC-alpha and its
CC   isoforms (including beta, gamma, delta, epsilon, zeta and eta) and so
CC   are useful for treating and diagnosing cell proliferation and
CC   differentiation processes such as neoplastic, hyperproliferative
```

```
CC   and inflammatory diseases.
CC   PNA oligomers have high affinity for complementary single stranded DNA.
CC   They are also able to form triple helices in which a first PNA strand
CC   binds with RNA or ssDNA and a second PNA strand binds with the resulting
CC   double helix or with the first PNA strand. The PNAs possess no
CC   significant charge and are water soluble, which facilitates cellular
CC   uptake. Further, since they contain amides of non-biological amino acids,
CC   they are biostable and resistant to enzymatic degradation by proteases.
CC   The present sequence targets the coding region of PKC-epsilon.
XX
SQ   Sequence 20 BP; 4 A; 11 C; 4 G; 1 T; 0 other;

  Query Match             66.0%;  Score 13.2;  DB 16;  Length 20;
  Best Local Similarity   83.3%;  Pred. No. 6.5e+03;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 CCACAGGGCTCACCAGGC 19
          || |||||| |||||| |
Db      1 CCCCAGGGCCCACCAGTC 18


RESULT 10
AAQ84237
ID   AAQ84237 standard; DNA; 20 BP.
XX
AC   AAQ84237;
XX
DT   21-SEP-1995  (first entry)
XX
DE   PKC-epsilon coding region antisense oligo, ISIS #7944.
XX
KW   Antisense; protein kinase C; alpha; PKC; beta; gamma; eta; epsilon;
KW   zeta; modulation; expression; isozyme; hybridise; 5' UTR; human;
KW   3' untranslated region; translation initiation site; detection;
KW   phosphorothioate linkage; 2'-O-methyl modification;
KW   2'-O-propyl modification; ss.
XX
OS   Synthetic.
XX
PN   WO9502069-A.
XX
PD   19-JAN-1995.
XX
PF   08-JUL-1994;   94WO-US07770.
XX
PR   09-JUL-1993;   93US-0089996.
PR   22-FEB-1994;   94US-0199779.
XX
PA   (ISIS-) ISIS PHARM INC.
XX
PI   Bennett CF,  Boggs RT,  Dean NM;
XX
DR   WPI; 1995-066911/09.
XX
PT   Oligo:nucleotide(s) hybridisable with Protein Kinase C mRNA or
PT   gene  - also novel PKC-alpha 3'-UTR sequence, useful for
PT   diagnosis and treatment of hyperproliferative disorders.
XX
PS   Claim 115; Page 37; 125pp; English.
XX
CC   The sequences given in AAQ84236-40 are oligos which are antisense to the
CC   protein kinase C-epsilon (PKC-epsilon) cDNA.  These antisense molecules
CC   may be used in modulating the expression of this particular isozyme of
CC   PKC.  The oligos of the invention preferably hybridise with the 5'- or
CC   3'-untranslated regions of the PKC gene, or the translation initiation
CC   site, or the coding region.  These oligos may be used in the detection
CC   of the human PKC genes and for treatment of animals with conditions
CC   associated with PKC, esp. hyperproliferative diseases such as psoriasis,
CC   colorectal cancer, lung cancer, breast or skin cancer.  These oligos may
CC   contain at least one phosphorothioate linkage and/or at least one of the
CC   nucleotides comprises a modification on the 2' position of the sugar,
CC   esp. a 2'-O-methyl or a 2'-O-propyl modification.
XX
SQ   Sequence 20 BP; 4 A; 11 C; 4 G; 1 T; 0 other;

  Query Match             66.0%;  Score 13.2;  DB 16;  Length 20;
  Best Local Similarity   83.3%;  Pred. No. 6.5e+03;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 CCACAGGGCTCACCAGGC 19
          || |||||| |||||| |
Db      1 CCCCAGGGCCCACCAGTC 18


RESULT 11
AAT44474
ID   AAT44474 standard; DNA; 20 BP.
XX
AC   AAT44474;
XX
DT   04-MAR-1997  (first entry)
XX
DE   Tick-borne encephalitis virus PCR primer T45.
XX
KW   Tick-borne encephalitis virus; TBEV; vaccine; virulence;
KW   replication; ss.
XX
OS   Synthetic.
XX
PN   WO9630521-A2.
XX
PD   03-OCT-1996.
XX
PF   28-MAR-1996;   96WO-EP01376.
XX
PR   31-MAR-1995;   95DE-4012142.
XX
PA   (IMMO ) IMMUNO AG.
XX
PI   Heinz FX,  Kunz C,  Mandl C;
XX
DR   WPI; 1996-455366/45.
XX
PT   cDNA encoding tick-borne encephalitis virus RNA - useful for vaccine
PT   prodn.
XX
PS   Example 1; Table 1 (Figure page 23/35); 126pp; German.
XX
CC   Example 1 describes the isolation of TBEV cDNA using the PCR
CC   method.  Primers are given in AAT44471 to AAT44485.  The isolated
CC   TBEV sequence is given in AAT44469.
XX
SQ   Sequence 20 BP; 6 A; 9 C; 4 G; 1 T; 0 other;

  Query Match             66.0%;  Score 13.2;  DB 17;  Length 20;
  Best Local Similarity   83.3%;  Pred. No. 6.5e+03;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      1 ACCACAGGGCTCACCAGG 18
          ||||||  || |||||||
Db      3 ACCACATAGCGCACCAGG 20


RESULT 12
AAT15828
ID   AAT15828 standard; DNA; 20 BP.
XX
AC   AAT15828;
XX
DT   07-AUG-1996  (first entry)
XX
DE   TBE virus strain Neudorfl genome oligonucleotide T45.
XX
KW   TBE virus; tick-borne encephalitis virus; Neudorfl;
KW   viral particle; vaccine; replication; ss.
XX
```

OS Synthetic.
XX
PN DE4426622-C1.
XX
PD 22-FEB-1996.
XX
PF 27-JUL-1994; 94DE-4426622.
XX
PR 27-JUL-1994; 94DE-4426622.
XX
PA (IMMO) IMMUNO AG.
XX
PI Heinz FX, Kunz C, Mandl C;
XX
DR WPI; 1996-106416/12.
XX
PT Complete DNA encoding tick borne encephalitis virus RNA - able to
PT produce infectious viral particles for use in vaccines
XX
PS Disclosure; Page 22; 33pp; German.
XX
CC The new complete DNA encoding tick borne encephalitis virus
CC RNA may be used to produce infectious viral particles. These
CC particles may be use in vaccines (live attenuated or inactivated),
CC opt. formulated with an immunomodulator. These virus particles are
CC also useful as seed for viral replication. The new cDNA is also
CC useful itself as a vaccine.
CC The sequence is the first complete TBEV sequence known and makes
CC possible targeted alteration of the viral genome (partic. for
CC attenuation or increasing viral yields). The nucleotides
CC used in the isolation of the sequence are given in AAT15825
CC to AAT15839.
XX
SQ Sequence 20 BP; 6 A; 9 C; 4 G; 1 T; 0 other;

Query Match 66.0%; Score 13.2; DB 17; Length 20;
Best Local Similarity 83.3%; Pred. No. 6.5e+03;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 ACCACAGGGCTCACCAGG 18
|||||| || |||||||
Db 3 ACCACATAGCGCACCAGG 20

RESULT 13
AAZ27354
ID AAZ27354 standard; DNA; 20 BP.
XX
AC AAZ27354;
XX
DT 01-DEC-1999 (first entry)
XX
DE Human protein kinase C epsilon antisense oligonucleotide #12.
XX
KW Human; protein kinase C; PKC; diagnosis; antisense oligonucleotide;
KW phosphorothioate; hybridisation; isozyme; target; inflammation;
KW hyperproliferative disorder; psoriasis; tumour; cancer; glioblastoma; ss.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN US5959096-A.
XX
PD 28-SEP-1999.
XX
PF 07-JUN-1995; 95US-0481066.
XX
PR 16-MAR-1992; 92US-0852852.
PR 09-JUL-1993; 93US-0089996.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Bennett CF, Dean N;
XX
DR WPI; 1999-561076/47.
XX
PT Antisense oligonucleotides useful for treatment of hyperproliferative
PT and inflammatory conditions including psoriasis, tumours and cancer -
XX
PS Example 16; Column 23; 56pp; English.
XX
CC The present invention describes antisense oligonucleotides up to 50
CC nucleotides in length which specifically bind mRNA encoding human
CC protein kinase C (PKC). AAZ27266 to AAZ27386 represent human PKC
CC antisense oligonucleotides used in the exemplification of the present
CC invention. The antisense oligonucleotides are useful for the treatment of
CC diseases associated with PKC expression, such as hyperproliferative and
CC inflammatory conditions including psoriasis, tumours and cancer
CC (glioblastoma, bladder, breast, colon and lung cancer).
XX
SQ Sequence 20 BP; 4 A; 11 C; 4 G; 1 T; 0 other;

Query Match 66.0%; Score 13.2; DB 20; Length 20;
Best Local Similarity 83.3%; Pred. No. 6.5e+03;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 CCACAGGGCTCACCAGGC 19
|| |||||| |||||| |
Db 1 CCCCAGGGCCCACCAGTC 18

RESULT 14
AAZ27368
ID AAZ27368 standard; DNA; 20 BP.
XX
AC AAZ27368;
XX
DT 01-DEC-1999 (first entry)
XX
DE Human protein kinase C epsilon antisense oligonucleotide #26.
XX
KW Human; protein kinase C; PKC; diagnosis; antisense oligonucleotide;
KW phosphorothioate; hybridisation; isozyme; target; inflammation;
KW hyperproliferative disorder; psoriasis; tumour; cancer; glioblastoma; ss.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN US5959096-A.
XX
PD 28-SEP-1999.
XX
PF 07-JUN-1995; 95US-0481066.
XX
PR 16-MAR-1992; 92US-0852852.
PR 09-JUL-1993; 93US-0089996.
XX
PA (ISIS-) ISIS PHARM INC.
XX
PI Bennett CF, Dean N;
XX
DR WPI; 1999-561076/47.
XX
PT Antisense oligonucleotides useful for treatment of hyperproliferative
PT and inflammatory conditions including psoriasis, tumours and cancer -
XX
PS Example 16; Column 23; 56pp; English.
XX
CC The present invention describes antisense oligonucleotides up to 50
CC nucleotides in length which specifically bind mRNA encoding human
CC protein kinase C (PKC). AAZ27266 to AAZ27386 represent human PKC
CC antisense oligonucleotides used in the exemplification of the present
CC invention. The antisense oligonucleotides are useful for the treatment of
CC diseases associated with PKC expression, such as hyperproliferative and
CC inflammatory conditions including psoriasis, tumours and cancer
CC (glioblastoma, bladder, breast, colon and lung cancer).

```
XX
SQ   Sequence 20 BP; 3 A; 12 C; 4 G; 1 T; 0 other;

  Query Match            66.0%;  Score 13.2;  DB 20;  Length 20;
  Best Local Similarity  83.3%;  Pred. No. 6.5e+03;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 CCACAGGGCTCACCAGGC 19
           || |||||| |||||| |
Db       2 CCCCAGGGCCCACCAGTC 19


RESULT 15
AAX78612
ID   AAX78612 standard; DNA; 20 BP.
XX
AC   AAX78612;
XX
DT   03-SEP-1999  (first entry)
XX
DE   Human PKC-epsilon oligonucleotide primer ISIS # 7944.
XX
KW   PKC; human; PKC-alpha; primer; protein kinase C; expression modulator;
KW   PKC-beta type I; PKC-beta type II; PKC-gamma; PKC-eta; PKC-delta;
KW   PKC-epsilon; PKC-zeta; anti-inflammatory; cytostatic;antisense targeting;
KW   isozyme; growth control; hyperproliferative disease; colon cancer;
KW   glioblastoma; bladder cancer; inflammatory condition; psoriasis; ss.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
PN   US5922686-A.
XX
PD   13-JUL-1999.
XX
PF   14-JUN-1996;   96US-0664336.
XX
PR   14-JUN-1996;   96US-0664336.
PR   16-MAR-1992;   92US-0852852.
PR   09-JUL-1993;   93US-0089996.
XX
PA   (ISIS-) ISIS PHARM INC.
XX
PI   Bennett CF,  Dean N;
XX
DR   WPI; 1999-404471/34.
XX
PT   Oligonucleotides targeted against nucleic acids encoding protein
PT   kinase C
XX
PS   Example 16; Column 63-64; 56pp; English.
XX
CC   This invention describes novel oligonucleotides (AAX78524-X78644) having
CC   up to 50 nucleotides hybridisable with, and able to modulate the
CC   expression of, a nucleic acid encoding protein kinase C and its isozymes
CC   alpha, beta type I, beta type II, gamma, eta, delta, epsilon and zeta.
CC   The oligonucleotides of the invention have anti-inflammatory and
CC   cytostatic activity and are used for antisense targeting to modulate the
CC   expression of PKC or of a particular PKC isozyme or set of isozymes in
CC   cells or tissues. The products of the invention also hybridise with
CC   nucleic acids involved in the modulation of PKC expression, which is
CC   known to be involved growth control in hyperproliferative diseases e.g.
CC   colon cancer, glioblastoma and bladder cancer as well as in inflammatory
CC   conditions e.g. psoriasis. Due to their specificity the oligonucleotides
CC   are able to overcome the problems of toxicity associated with previous
CC   agents designed to modulate PKC expression.
XX
SQ   Sequence 20 BP; 4 A; 11 C; 4 G; 1 T; 0 other;

  Query Match            66.0%;  Score 13.2;  DB 20;  Length 20;
  Best Local Similarity  83.3%;  Pred. No. 6.5e+03;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 CCACAGGGCTCACCAGGC 19
           || |||||| |||||| |
Db       1 CCCCAGGGCCCACCAGTC 18


Search completed: November 23, 2002, 07:04:07
Job time : 97.1 secs
```

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 06:36:31 ; Search time 21.3 Seconds
(without alignments)
287.959 Million cell updates/sec

Title: US-09-296-264-30
Perfect score: 20
Sequence: 1 accacagggctcaccaggcg 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 687286

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
c	1	14.2	71.0	93	4	US-09-461-697-33	Sequence 33, Appl
c	2	14.2	71.0	97	1	US-08-447-172A-19	Sequence 19, Appl
	3	13.2	66.0	20	2	US-08-478-178A-89	Sequence 89, Appl
	4	13.2	66.0	20	2	US-08-478-178A-103	Sequence 103, App
	5	13.2	66.0	20	2	US-08-488-177-89	Sequence 89, Appl
	6	13.2	66.0	20	2	US-08-488-177-103	Sequence 103, App
	7	13.2	66.0	20	2	US-08-481-072A-89	Sequence 89, Appl
	8	13.2	66.0	20	2	US-08-481-072A-103	Sequence 103, App
	9	13.2	66.0	20	2	US-08-664-336-89	Sequence 89, Appl
	10	13.2	66.0	20	2	US-08-664-336-103	Sequence 103, App
	11	13.2	66.0	20	2	US-08-481-066A-89	Sequence 89, Appl
	12	13.2	66.0	20	2	US-08-481-066A-103	Sequence 103, App
	13	13.2	66.0	20	3	US-08-578-615A-97	Sequence 97, Appl
	14	13.2	66.0	20	4	US-08-829-637A-89	Sequence 89, Appl
	15	13.2	66.0	20	4	US-08-829-637A-103	Sequence 103, App
	16	13.2	66.0	20	5	PCT-US94-07770-97	Sequence 97, Appl
c	17	13.2	66.0	37	1	US-08-292-926-1	Sequence 1, Appli
	18	13.2	66.0	38	1	US-08-373-124A-1181	Sequence 1181, Ap
	19	13.2	66.0	38	1	US-08-435-628-1181	Sequence 1181, Ap
c	20	13.2	66.0	50	2	US-08-299-074A-19	Sequence 19, Appl
	21	13.2	66.0	50	2	US-08-687-355A-8	Sequence 8, Appli
c	22	13.2	66.0	50	4	US-09-399-773-19	Sequence 19, Appl
	23	13.2	66.0	50	4	US-09-407-367-8	Sequence 8, Appli
c	24	13.2	66.0	60	6	5194596-31	Patent No. 5194596
c	25	13.2	66.0	60	6	5219739-36	Patent No. 5219739
c	26	13.2	66.0	96	1	US-08-292-926-2	Sequence 2, Appli
c	27	13	65.0	21	3	US-09-009-913-49	Sequence 49, Appl
c	28	12.8	64.0	24	4	US-08-891-292A-90	Sequence 90, Appl
c	29	12.6	63.0	24	2	US-08-691-814B-68	Sequence 68, Appl
	30	12.6	63.0	56	1	US-08-652-245-51	Sequence 51, Appl
	31	12.6	63.0	56	2	US-09-079-186-51	Sequence 51, Appl
	32	12.6	63.0	56	4	US-09-274-647-51	Sequence 51, Appl
	33	12.6	63.0	67	1	US-07-972-032-42	Sequence 42, Appl
	34	12.6	63.0	67	1	US-08-642-255-56	Sequence 56, Appl
c	35	12.6	63.0	75	1	US-07-972-032-43	Sequence 43, Appl
c	36	12.6	63.0	75	1	US-08-642-255-57	Sequence 57, Appl
	37	12.6	63.0	87	1	US-08-433-126A-44	Sequence 44, Appl
	38	12.6	63.0	87	1	US-08-433-124A-44	Sequence 44, Appl
	39	12.6	63.0	87	3	US-08-976-413A-44	Sequence 44, Appl
	40	12.6	63.0	87	5	PCT-US96-06059-44	Sequence 44, Appl
c	41	12.4	62.0	94	1	US-08-188-277B-21	Sequence 21, Appl
	42	12.2	61.0	22	4	US-08-943-731-517	Sequence 517, App
c	43	12.2	61.0	25	2	US-08-766-982-7	Sequence 7, Appli
c	44	12.2	61.0	25	4	US-09-296-219-7	Sequence 7, Appli
c	45	12.2	61.0	30	2	US-08-678-039A-24	Sequence 24, Appl

ALIGNMENTS

RESULT 1
US-09-461-697-33/c
; Sequence 33, Application US/09461697
; Patent No. 6277974
; GENERAL INFORMATION:
; APPLICANT: COGENT NEUROSCIENCE, Inc.
; APPLICANT: Lo, Donald C.
; APPLICANT: Barney, Shawn
; APPLICANT: Thomas, Mary Beth
; APPLICANT: Portbury, Stuart D.
; APPLICANT: Puranam, Kasturi
; APPLICANT: Katz, Lawrence C.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSING
; TITLE OF INVENTION: AND TREATING CONDITIONS, DISORDERS, OR DISEASES INVOLVING
; TITLE OF INVENTION: CELL DEATH
; FILE REFERENCE: 10001-005-999
; CURRENT APPLICATION NUMBER: US/09/461,697
; CURRENT FILING DATE: 1999-12-14
; NUMBER OF SEQ ID NOS: 466
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 33
; LENGTH: 93
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-461-697-33

Query Match 71.0%; Score 14.2; DB 4; Length 93;
Best Local Similarity 84.2%; Pred. No. 3.5e+02;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy       2 CCACAGGGCTCACCAGGCG 20
           |||| |||||| |||| ||
Db      28 CCACTGGGCTCCCCAGTCG 10
```

RESULT 2
US-08-447-172A-19/c
; Sequence 19, Application US/08447172A
; Patent No. 5726017
; GENERAL INFORMATION:
; APPLICANT: MICHAEL LOCHRIE AND LARRY GOLD
; TITLE OF INVENTION: HIGH AFFINITY HIV-1 GAG
; TITLE OF INVENTION: NUCLEIC ACID LIGANDS
; NUMBER OF SEQUENCES: 47
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Swanson and Bratschun, L.L.C.
; STREET: 8400 East Prentice Avenue, Suite #200
; CITY: Denver
; STATE: Colorado
; COUNTRY: USA

```
;      ZIP:  80111
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.50 inch, 1.44 Mb storage
;      COMPUTER:  IBM compatible
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  WordPerfect 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/447,172A
;      FILING DATE:  19-MAY- 1995
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/714,131
;      FILING DATE:  10-JUNE-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/931,473
;      FILING DATE:  17-AUGUST-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/964,624
;      FILING DATE:  21-OCTOBER-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/117,991
;      FILING DATE:  8-SEPTEMBER-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/536,428
;      FILING DATE:  11-JUNE-1990
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Diane H. McClearn
;      REGISTRATION NUMBER:  33,960
;      REFERENCE/DOCKET NUMBER:  NEX32
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (303) 793-3333
;      TELEFAX:  (303) 793-3433
;  INFORMATION FOR SEQ ID NO:  19:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  97 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  RNA
US-08-447-172A-19

  Query Match             71.0%;  Score 14.2;  DB 1;  Length 97;
  Best Local Similarity   84.2%;  Pred. No. 3.5e+02;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 CCACAGGGCTCACCAGGCG 20
           | ||||||||||||  ||
Db      44 CGACAGGGCTCACCATCCG 26


RESULT 3
US-08-478-178A-89
; Sequence 89, Application US/08478178A
; Patent No. 5882927
;  GENERAL INFORMATION:
;    APPLICANT:  Nicholas Dean, C. Frank Bennett
;    TITLE OF INVENTION:  Oligonucleotide Modulation of
;    TITLE OF INVENTION:  Protein                                  Kinase C
;    NUMBER OF SEQUENCES:  121
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Woodcock Washburn Kurtz
;      ADDRESSEE:  Mackiewicz & No. 5882927ris
;      STREET:  One Liberty Place - 46th Floor
;      CITY:  Philadelphia
;      STATE:  PA
;      COUNTRY:  USA
;      ZIP:  19103
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  DISKETTE, 3.5 INCH, 1.44 Mb STORAGE
;      COMPUTER:  IBM PS/2
;      OPERATING SYSTEM:  PC-DOS
;      SOFTWARE:  WORDPERFECT 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/478,178A
;      FILING DATE:  herewith
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  852,852
;      FILING DATE:  March 16, 1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Rebecca Ralph Gaumond
;      REGISTRATION NUMBER:  35,152
;      REFERENCE/DOCKET NUMBER:  ISIS-1154
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (215) 568-3100
;      TELEFAX:  (215) 568-3439
;  INFORMATION FOR SEQ ID NO:  89:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    ANTI-SENSE:  yes
US-08-478-178A-89

  Query Match             66.0%;  Score 13.2;  DB 2;  Length 20;
  Best Local Similarity   83.3%;  Pred. No. 9.8e+02;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 CCACAGGGCTCACCAGGC 19
           || |||||| |||||| |
Db       1 CCCCAGGGCCCACCAGTC 18


RESULT 4
US-08-478-178A-103
; Sequence 103, Application US/08478178A
; Patent No. 5882927
;  GENERAL INFORMATION:
;    APPLICANT:  Nicholas Dean, C. Frank Bennett
;    TITLE OF INVENTION:  Oligonucleotide Modulation of
;    TITLE OF INVENTION:  Protein                                  Kinase C
;    NUMBER OF SEQUENCES:  121
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Woodcock Washburn Kurtz
;      ADDRESSEE:  Mackiewicz & No. 5882927ris
;      STREET:  One Liberty Place - 46th Floor
;      CITY:  Philadelphia
;      STATE:  PA
;      COUNTRY:  USA
;      ZIP:  19103
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  DISKETTE, 3.5 INCH, 1.44 Mb STORAGE
;      COMPUTER:  IBM PS/2
;      OPERATING SYSTEM:  PC-DOS
;      SOFTWARE:  WORDPERFECT 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/478,178A
;      FILING DATE:  herewith
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  852,852
;      FILING DATE:  March 16, 1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Rebecca Ralph Gaumond
;      REGISTRATION NUMBER:  35,152
;      REFERENCE/DOCKET NUMBER:  ISIS-1154
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (215) 568-3100
;      TELEFAX:  (215) 568-3439
;  INFORMATION FOR SEQ ID NO:  103:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
```

```
;    ANTI-SENSE:  yes
US-08-478-178A-103

  Query Match             66.0%;  Score 13.2;  DB 2;  Length 20;
  Best Local Similarity   83.3%;  Pred. No. 9.8e+02;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 CCACAGGGCTCACCAGGC 19
          || |||||| |||||| |
Db      2 CCCCAGGGCCCACCAGTC 19


RESULT 5
US-08-488-177-89
; Sequence 89, Application US/08488177
; Patent No. 5885970
;  GENERAL INFORMATION:
;    APPLICANT:  Nicholas Dean, C. Frank Bennett
;    TITLE OF INVENTION:  Oligonucleotide Modulation of
;    TITLE OF INVENTION:  Protein Kinase C
;    NUMBER OF SEQUENCES:  121
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Woodcock Washburn Kurtz
;      ADDRESSEE:  Mackiewicz & No. 5885970ris
;      STREET:  One Liberty Place - 46th Floor
;      CITY:  Philadelphia
;      STATE:  PA
;      COUNTRY:  USA
;      ZIP:  19103
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  DISKETTE, 3.5 INCH, 1.44 Mb STORAGE
;      COMPUTER:  IBM PS/2
;      OPERATING SYSTEM:  PC-DOS
;      SOFTWARE:  WORDPERFECT 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/488,177
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  852,852
;      FILING DATE:  March 16, 1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Paul K. Legaard
;      REGISTRATION NUMBER:  38,534
;      REFERENCE/DOCKET NUMBER:  ISIS-1995
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (215) 568-3100
;      TELEFAX:  (215) 568-3439
;  INFORMATION FOR SEQ ID NO:  89:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    ANTI-SENSE:  yes
US-08-488-177-89

  Query Match             66.0%;  Score 13.2;  DB 2;  Length 20;
  Best Local Similarity   83.3%;  Pred. No. 9.8e+02;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 CCACAGGGCTCACCAGGC 19
          || |||||| |||||| |
Db      1 CCCCAGGGCCCACCAGTC 18


RESULT 6
US-08-488-177-103
; Sequence 103, Application US/08488177
; Patent No. 5885970
;  GENERAL INFORMATION:
;    APPLICANT:  Nicholas Dean, C. Frank Bennett
;    TITLE OF INVENTION:  Oligonucleotide Modulation of
;    TITLE OF INVENTION:  Protein Kinase C
;    NUMBER OF SEQUENCES:  121
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Woodcock Washburn Kurtz
;      ADDRESSEE:  Mackiewicz & No. 5885970ris
;      STREET:  One Liberty Place - 46th Floor
;      CITY:  Philadelphia
;      STATE:  PA
;      COUNTRY:  USA
;      ZIP:  19103
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  DISKETTE, 3.5 INCH, 1.44 Mb STORAGE
;      COMPUTER:  IBM PS/2
;      OPERATING SYSTEM:  PC-DOS
;      SOFTWARE:  WORDPERFECT 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/488,177
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  852,852
;      FILING DATE:  March 16, 1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Paul K. Legaard
;      REGISTRATION NUMBER:  38,534
;      REFERENCE/DOCKET NUMBER:  ISIS-1995
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (215) 568-3100
;      TELEFAX:  (215) 568-3439.
;  INFORMATION FOR SEQ ID NO:  103:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    ANTI-SENSE:  yes
US-08-488-177-103

  Query Match             66.0%;  Score 13.2;  DB 2;  Length 20;
  Best Local Similarity   83.3%;  Pred. No. 9.8e+02;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 CCACAGGGCTCACCAGGC 19
          || |||||| |||||| |
Db      2 CCCCAGGGCCCACCAGTC 19


RESULT 7
US-08-481-072A-89
; Sequence 89, Application US/08481072A
; Patent No. 5916807
;  GENERAL INFORMATION:
;    APPLICANT:  Nicholas Dean, C. Frank Bennett
;    TITLE OF INVENTION:  Oligonucleotide Modulation of
;    TITLE OF INVENTION:  Protein                                    Kinase C
;    NUMBER OF SEQUENCES:  121
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Woodcock Washburn Kurtz
;      ADDRESSEE:  Mackiewicz & No. 5916807ris
;      STREET:  One Liberty Place - 46th Floor
;      CITY:  Philadelphia
;      STATE:  PA
;      COUNTRY:  USA
;      ZIP:  19103
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  DISKETTE, 3.5 INCH, 1.44 Mb STORAGE
;      COMPUTER:  IBM PS/2
;      OPERATING SYSTEM:  PC-DOS
;      SOFTWARE:  WORDPERFECT 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/481,072A
;      FILING DATE:  herewith
;      CLASSIFICATION:  514
```

```
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  852,852
;      FILING DATE:  March 16, 1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Rebecca Ralph Gaumond
;      REGISTRATION NUMBER:  35,152
;      REFERENCE/DOCKET NUMBER:  ISIS-1154
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (215) 568-3100
;      TELEFAX:  (215) 568-3439
;  INFORMATION FOR SEQ ID NO:  89:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    ANTI-SENSE:  yes
US-08-481-072A-89

  Query Match             66.0%;  Score 13.2;  DB 2;  Length 20;
  Best Local Similarity   83.3%;  Pred. No. 9.8e+02;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 CCACAGGGCTCACCAGGC 19
          || |||||| |||||| |
Db      1 CCCCAGGGCCCACCAGTC 18


RESULT 8
US-08-481-072A-103
; Sequence 103, Application US/08481072A
; Patent No. 5916807
;  GENERAL INFORMATION:
;    APPLICANT:  Nicholas Dean, C. Frank Bennett
;    TITLE OF INVENTION:  Oligonucleotide Modulation of
;    TITLE OF INVENTION:  Protein                                Kinase C
;    NUMBER OF SEQUENCES:  121
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Woodcock Washburn Kurtz
;      ADDRESSEE:  Mackiewicz & No. 5916807ris
;      STREET:  One Liberty Place - 46th Floor
;      CITY:  Philadelphia
;      STATE:  PA
;      COUNTRY:  USA
;      ZIP:  19103
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  DISKETTE, 3.5 INCH, 1.44 Mb STORAGE
;      COMPUTER:  IBM PS/2
;      OPERATING SYSTEM:  PC-DOS
;      SOFTWARE:  WORDPERFECT 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/481,072A
;      FILING DATE:  herewith
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  852,852
;      FILING DATE:  March 16, 1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Rebecca Ralph Gaumond
;      REGISTRATION NUMBER:  35,152
;      REFERENCE/DOCKET NUMBER:  ISIS-1154
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (215) 568-3100
;      TELEFAX:  (215) 568-3439
;  INFORMATION FOR SEQ ID NO:  103:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    ANTI-SENSE:  yes
US-08-481-072A-103

  Query Match             66.0%;  Score 13.2;  DB 2;  Length 20;
  Best Local Similarity   83.3%;  Pred. No. 9.8e+02;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 CCACAGGGCTCACCAGGC 19
          || |||||| |||||| |
Db      2 CCCCAGGGCCCACCAGTC 19


RESULT 9
US-08-664-336-89
; Sequence 89, Application US/08664336
; Patent No. 5922686
;  GENERAL INFORMATION:
;    APPLICANT:  Nicholas Dean, C. Frank Bennett
;    TITLE OF INVENTION:  Oligonucleotide Modulation of Protein
;    NUMBER OF SEQUENCES:  121
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Woodcock Washburn Kurtz
;      ADDRESSEE:  Mackiewicz & No. 5922686ris
;      STREET:  One Liberty Place - 46th Floor
;      CITY:  Philadelphia
;      STATE:  PA
;      COUNTRY:  USA
;      ZIP:  19103
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  DISKETTE, 3.5 INCH, 720 kb STORAGE
;      COMPUTER:  IBM PS/2
;      OPERATING SYSTEM:  PC-DOS
;      SOFTWARE:  WORDPERFECT 6.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/664,336
;      FILING DATE:  herewith
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  852,852
;      FILING DATE:  March 16, 1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  089,996
;      FILING DATE:  July 9, 1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Paul K. Legaard
;      REGISTRATION NUMBER:  38,534
;      REFERENCE/DOCKET NUMBER:  ISIS-2345
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (215) 568-3100
;      TELEFAX:  (215) 568-3439
;  INFORMATION FOR SEQ ID NO:  89:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    ANTI-SENSE:  yes
US-08-664-336-89

  Query Match             66.0%;  Score 13.2;  DB 2;  Length 20;
  Best Local Similarity   83.3%;  Pred. No. 9.8e+02;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 CCACAGGGCTCACCAGGC 19
          || |||||| |||||| |
Db      1 CCCCAGGGCCCACCAGTC 18


RESULT 10
US-08-664-336-103
; Sequence 103, Application US/08664336
; Patent No. 5922686
;  GENERAL INFORMATION:
;    APPLICANT:  Nicholas Dean, C. Frank Bennett
;    TITLE OF INVENTION:  Oligonucleotide Modulation of Protein
;    NUMBER OF SEQUENCES:  121
```

```
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Woodcock Washburn Kurtz
;      ADDRESSEE:  Mackiewicz & No. 5922686ris
;      STREET:  One Liberty Place - 46th Floor
;      CITY:  Philadelphia
;      STATE:  PA
;      COUNTRY:  USA
;      ZIP:  19103
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  DISKETTE, 3.5 INCH, 720 kb STORAGE
;      COMPUTER:  IBM PS/2
;      OPERATING SYSTEM:  PC-DOS
;      SOFTWARE:  WORDPERFECT 6.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/664,336
;      FILING DATE:  herewith
;      CLASSIFICATION:  536
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  852,852
;      FILING DATE:  March 16, 1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  089,996
;      FILING DATE:  July 9, 1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Paul K. Legaard
;      REGISTRATION NUMBER:  38,534
;      REFERENCE/DOCKET NUMBER:  ISIS-2345
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (215) 568-3100
;      TELEFAX:  (215) 568-3439
;  INFORMATION FOR SEQ ID NO:  103:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    ANTI-SENSE:  yes
US-08-664-336-103

  Query Match             66.0%;  Score 13.2;  DB 2;  Length 20;
  Best Local Similarity   83.3%;  Pred. No. 9.8e+02;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 CCACAGGGCTCACCAGGC 19
           || |||||| |||||| |
Db       2 CCCCAGGGCCCACCAGTC 19


RESULT 11
US-08-481-066A-89
; Sequence 89, Application US/08481066A
; Patent No. 5959096
;  GENERAL INFORMATION:
;    APPLICANT:  Nicholas Dean, C. Frank Bennett
;    TITLE OF INVENTION:  Oligonucleotide Modulation of
;    TITLE OF INVENTION:  Protein Kinase C
;    NUMBER OF SEQUENCES:  121
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Woodcock Washburn Kurtz
;      ADDRESSEE:  Mackiewicz & No. 5959096ris
;      STREET:  One Liberty Place - 46th Floor
;      CITY:  Philadelphia
;      STATE:  PA
;      COUNTRY:  USA
;      ZIP:  19103
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  DISKETTE, 3.5 INCH, 1.44 Mb STORAGE
;      COMPUTER:  IBM PS/2
;      OPERATING SYSTEM:  PC-DOS
;      SOFTWARE:  WORDPERFECT 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/481,066A
;      FILING DATE:  herewith
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  852,852
;      FILING DATE:  March 16, 1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Rebecca Ralph Gaumond
;      REGISTRATION NUMBER:  35,152
;      REFERENCE/DOCKET NUMBER:  ISIS-1154
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (215) 568-3100
;      TELEFAX:  (215) 568-3439
;  INFORMATION FOR SEQ ID NO:  89:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    ANTI-SENSE:  yes
US-08-481-066A-89

  Query Match             66.0%;  Score 13.2;  DB 2;  Length 20;
  Best Local Similarity   83.3%;  Pred. No. 9.8e+02;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 CCACAGGGCTCACCAGGC 19
           || |||||| |||||| |
Db       1 CCCCAGGGCCCACCAGTC 18


RESULT 12
US-08-481-066A-103
; Sequence 103, Application US/08481066A
; Patent No. 5959096
;  GENERAL INFORMATION:
;    APPLICANT:  Nicholas Dean, C. Frank Bennett
;    TITLE OF INVENTION:  Oligonucleotide Modulation of
;    TITLE OF INVENTION:  Protein Kinase C
;    NUMBER OF SEQUENCES:  121
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Woodcock Washburn Kurtz
;      ADDRESSEE:  Mackiewicz & No. 5959096ris
;      STREET:  One Liberty Place - 46th Floor
;      CITY:  Philadelphia
;      STATE:  PA
;      COUNTRY:  USA
;      ZIP:  19103
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  DISKETTE, 3.5 INCH, 1.44 Mb STORAGE
;      COMPUTER:  IBM PS/2
;      OPERATING SYSTEM:  PC-DOS
;      SOFTWARE:  WORDPERFECT 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/481,066A
;      FILING DATE:  herewith
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  852,852
;      FILING DATE:  March 16, 1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Rebecca Ralph Gaumond
;      REGISTRATION NUMBER:  35,152
;      REFERENCE/DOCKET NUMBER:  ISIS-1154
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (215) 568-3100
;      TELEFAX:  (215) 568-3439
;  INFORMATION FOR SEQ ID NO:  103:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    ANTI-SENSE:  yes
US-08-481-066A-103
```

```
  Query Match             66.0%;  Score 13.2;  DB 2;  Length 20;
  Best Local Similarity   83.3%;  Pred. No. 9.8e+02;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 CCACAGGGCTCACCAGGC 19
          || |||||| |||||| |
Db      2 CCCCAGGGCCCACCAGTC 19


RESULT 13
US-08-578-615A-97
; Sequence 97, Application US/08578615A
; Patent No. 6015892
;  GENERAL INFORMATION:
;    APPLICANT:  Nicholas Dean, C. Frank Bennett and Russell, T. Boggs
;    TITLE OF INVENTION:  Oligonucleotide Modulation of Protein KinaseC
;    NUMBER OF SEQUENCES:  122
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Woodcock Washburn Kurtz Mackiewicz & No. 6015892ris LLP
;      STREET:  One Liberty Place - 46th Floor
;      CITY:  Philadelphia
;      STATE:  PA
;      COUNTRY:  USA
;      ZIP:  19103
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  DISKETTE, 3.5 INCH, 1.44 Mb STORAGE
;      COMPUTER:  IBM PS/2
;      OPERATING SYSTEM:  PC-DOS
;      SOFTWARE:  WORDPERFECT 6.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/578,615A
;      FILING DATE:  11-JAN-1996
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  852,852
;      FILING DATE:  16-MAR-1992
;      APPLICATION NUMBER:  08/089,996
;      FILING DATE:  09-JUL-1993
;      APPLICATION NUMBER:  08/199,779
;      FILING DATE:  22-FEB-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Paul K. Legaard
;      REGISTRATION NUMBER:  38,534
;      REFERENCE/DOCKET NUMBER:  ISIS-1568
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (215) 568-3100
;      TELEFAX:  (215) 568-3439
;  INFORMATION FOR SEQ ID NO:  97:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    ANTI-SENSE:  yes
US-08-578-615A-97

  Query Match             66.0%;  Score 13.2;  DB 3;  Length 20;
  Best Local Similarity   83.3%;  Pred. No. 9.8e+02;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 CCACAGGGCTCACCAGGC 19
          || |||||| |||||| |
Db      1 CCCCAGGGCCCACCAGTC 18


RESULT 14
US-08-829-637A-89
; Sequence 89, Application US/08829637A
; Patent No. 6339066
;  GENERAL INFORMATION:
;    APPLICANT:  C. Frank Bennett
;    APPLICANT:  Phillip Dan Cook
;    APPLICANT:  Nicholas Dean
;    APPLICANT:  Glenn Hoke
;    TITLE OF INVENTION:  OLIGONUCLEOTIDES WHICH HAVE
;    TITLE OF INVENTION:  PHOSPHOROTHIOATE LINKAGES OF HIGH CHIRAL PURITY AND
;    TITLE OF INVENTION:  WHICH MODULATE aI, aII,  , k, n,   AND   ISOFORMS OF
;    TITLE OF INVENTION:  PROTEIN KINASE C
;    NUMBER OF SEQUENCES:  136
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  John W. Caldwell (28,937) Woodcock
;      ADDRESSEE:  Washburn Kurtz Mackiewicz & No. 6339066ris
;      STREET:  One Liberty Place - 46th Floor
;      CITY:  Philadelphia
;      STATE:  PA
;      COUNTRY:  USA
;      ZIP:  19103
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  DISKETTE, 3.5 INCH, 1.44 Mb STORAGE
;      COMPUTER:  IBM PS/2
;      OPERATING SYSTEM:  PC-DOS
;      SOFTWARE:  WORDPERFECT 6.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/829,637A
;      FILING DATE:  herewith
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/481,066
;      FILING DATE:  07-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/470,129
;      FILING DATE:  06-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/469,851
;      FILING DATE:  06-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/468,569
;      FILING DATE:  06-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/089,996
;      FILING DATE:  09-JUL-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/058,023
;      FILING DATE:  05-MAY-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/777,007
;      FILING DATE:  16-OCT-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/777,760
;      FILING DATE:  15-OCT-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/852,852
;      FILING DATE:  16-MAR-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US91/00243
;      FILING DATE:  11-JAN-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/566,977
;      FILING DATE:  13-AUG-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/436,358
;      FILING DATE:  11-JAN-1990
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  _____
;      REGISTRATION NUMBER:  _____
;      REFERENCE/DOCKET NUMBER:  ISIS-____
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (215) 568-3100
;      TELEFAX:  (215) 568-3439
;  INFORMATION FOR SEQ ID NO:  89:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
```

```
;    ANTI-SENSE:  yes
US-08-829-637A-89

  Query Match             66.0%;  Score 13.2;  DB 4;  Length 20;
  Best Local Similarity   83.3%;  Pred. No. 9.8e+02;
  Matches   15;  Conservative    0;  Mismatches   3;  Indels    0;  Gaps    0;

Qy      2 CCACAGGGCTCACCAGGC 19
          || |||||| |||||| |
Db      1 CCCCAGGGCCCACCAGTC 18


RESULT 15
US-08-829-637A-103
; Sequence 103, Application US/08829637A
; Patent No. 6339066
;  GENERAL INFORMATION:
;    APPLICANT:  C. Frank Bennett
;    APPLICANT:  Phillip Dan Cook
;    APPLICANT:  Nicholas Dean
;    APPLICANT:  Glenn Hoke
;    TITLE OF INVENTION:  OLIGONUCLEOTIDES WHICH HAVE
;    TITLE OF INVENTION:  PHOSPHOROTHIOATE LINKAGES OF HIGH CHIRAL PURITY AND
;    TITLE OF INVENTION:  WHICH MODULATE aI, aII,  , k, n,   AND   ISOFORMS OF
;    TITLE OF INVENTION:  PROTEIN KINASE C
;    NUMBER OF SEQUENCES:  136
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  John W. Caldwell (28,937) Woodcock
;      ADDRESSEE:  Washburn Kurtz Mackiewicz & No. 6339066ris
;      STREET:  One Liberty Place - 46th Floor
;      CITY:  Philadelphia
;      STATE:  PA
;      COUNTRY:  USA
;      ZIP:  19103
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  DISKETTE, 3.5 INCH, 1.44 Mb STORAGE
;      COMPUTER:  IBM PS/2
;      OPERATING SYSTEM:  PC-DOS
;      SOFTWARE:  WORDPERFECT 6.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/829,637A
;      FILING DATE:  herewith
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/481,066
;      FILING DATE:  07-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/470,129
;      FILING DATE:  06-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/469,851
;      FILING DATE:  06-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/468,569
;      FILING DATE:  06-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/089,996
;      FILING DATE:  09-JUL-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/058,023
;      FILING DATE:  05-MAY-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/777,007
;      FILING DATE:  16-OCT-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/777,760
;      FILING DATE:  15-OCT-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/852,852
;      FILING DATE:  16-MAR-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US91/00243
;      FILING DATE:  11-JAN-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/566,977
;      FILING DATE:  13-AUG-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/436,358
;      FILING DATE:  11-JAN-1990
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  _____
;      REGISTRATION NUMBER:  _____
;      REFERENCE/DOCKET NUMBER:  ISIS-____
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (215) 568-3100
;      TELEFAX:  (215) 568-3439
;  INFORMATION FOR SEQ ID NO:  103:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  20
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    ANTI-SENSE:  yes
US-08-829-637A-103

  Query Match             66.0%;  Score 13.2;  DB 4;  Length 20;
  Best Local Similarity   83.3%;  Pred. No. 9.8e+02;
  Matches   15;  Conservative    0;  Mismatches   3;  Indels    0;  Gaps    0;

Qy      2 CCACAGGGCTCACCAGGC 19
          || |||||| |||||| |
Db      2 CCCCAGGGCCCACCAGTC 19


Search completed: November 23, 2002, 07:07:45
Job time : 23.3 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 23, 2002, 06:42:25 ; Search time 16.8 Seconds
(without alignments)
450.869 Million cell updates/sec

Title: US-09-296-264-30
Perfect score: 20
Sequence: 1 accacagggctcaccaggcg 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 335578 seqs, 189365133 residues

Total number of hits satisfying chosen parameters: 195614

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	14.2	71.0	93	10	US-09-922-261-33	Sequence 33, Appl
c	2	13.2	66.0	50	9	US-09-813-824A-19	Sequence 19, Appl
	3	13.2	66.0	90	10	US-09-864-761-20746	Sequence 20746, A
	4	12.8	64.0	24	10	US-09-761-962-37	Sequence 37, Appl
	5	12.8	64.0	31	10	US-09-801-274-874	Sequence 874, App
c	6	12.8	64.0	79	10	US-09-783-590-11990	Sequence 11990, A
	7	12.6	63.0	56	10	US-09-759-136-51	Sequence 51, Appl
	8	12.4	62.0	31	10	US-09-801-274-1365	Sequence 1365, Ap
	9	12.2	61.0	36	10	US-09-504-231A-2132	Sequence 2132, Ap
	10	12.2	61.0	36	10	US-09-274-553D-2132	Sequence 2132, Ap
	11	12.2	61.0	39	10	US-09-484-704-56	Sequence 56, Appl
c	12	12.2	61.0	75	10	US-09-864-761-29557	Sequence 29557, A
c	13	12	60.0	97	10	US-09-864-761-24315	Sequence 24315, A
	14	11.8	59.0	22	10	US-09-897-438B-6	Sequence 6, Appli
	15	11.8	59.0	31	10	US-09-801-274-915	Sequence 915, App
c	16	11.8	59.0	88	10	US-09-783-590-11987	Sequence 11987, A
	17	11.8	59.0	91	10	US-09-764-877-3369	Sequence 3369, Ap
	18	11.8	59.0	96	10	US-09-294-093B-4238	Sequence 4238, Ap
	19	11.8	59.0	98	10	US-09-294-093B-2481	Sequence 2481, Ap
	20	11.8	59.0	100	10	US-09-999-672-20	Sequence 20, Appl
	21	11.8	59.0	100	12	US-10-040-863-20	Sequence 20, Appl
	22	11.6	58.0	22	10	US-09-875-338-27	Sequence 27, Appl
c	23	11.6	58.0	27	10	US-09-817-014-85	Sequence 85, Appl
c	24	11.6	58.0	79	9	US-10-079-623-198	Sequence 198, App
	25	11.6	58.0	87	10	US-09-191-724-3	Sequence 3, Appli
c	26	11.6	58.0	88	10	US-09-864-761-21663	Sequence 21663, A
c	27	11.4	57.0	20	10	US-09-752-639-64	Sequence 64, Appl
c	28	11.4	57.0	20	10	US-09-984-198-64	Sequence 64, Appl
	29	11.4	57.0	21	9	US-09-988-113-21	Sequence 21, Appl
c	30	11.4	57.0	21	10	US-09-765-081-185	Sequence 185, App
	31	11.4	57.0	21	10	US-09-776-874A-21	Sequence 21, Appl
	32	11.4	57.0	23	9	US-09-905-291A-400	Sequence 400, App
c	33	11.4	57.0	23	9	US-09-976-736-26	Sequence 26, Appl
c	34	11.4	57.0	23	9	US-09-976-736-28	Sequence 28, Appl
c	35	11.4	57.0	23	9	US-09-988-113-24	Sequence 24, Appl
c	36	11.4	57.0	23	10	US-09-776-874A-24	Sequence 24, Appl
	37	11.4	57.0	23	10	US-09-909-320-400	Sequence 400, App
	38	11.4	57.0	23	10	US-09-909-088B-400	Sequence 400, App
c	39	11.4	57.0	31	10	US-09-801-274-420	Sequence 420, App
	40	11.4	57.0	31	10	US-09-801-274-882	Sequence 882, App
	41	11.4	57.0	69	10	US-09-833-381-1859	Sequence 1859, Ap
	42	11.4	57.0	79	10	US-09-864-761-24159	Sequence 24159, A
c	43	11.2	56.0	17	10	US-09-263-959-4	Sequence 4, Appli
c	44	11.2	56.0	18	10	US-09-863-179-15	Sequence 15, Appl
c	45	11.2	56.0	20	10	US-09-815-068A-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-09-922-261-33/c
; Sequence 33, Application US/09922261
; Patent No. US20020111471A1
; GENERAL INFORMATION:
; APPLICANT: COGENT NEUROSCIENCE, Inc.
; APPLICANT: Lo, Donald C.
; APPLICANT: Barney, Shawn
; APPLICANT: Thomas, Mary Beth
; APPLICANT: Portbury, Stuart D.
; APPLICANT: Puranam, Kasturi
; APPLICANT: Katz, Lawrence C.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSING
; TITLE OF INVENTION: AND TREATING CONDITIONS, DISORDERS, OR DISEASES INVOLVING
; TITLE OF INVENTION: CELL DEATH
; FILE REFERENCE: 10001-005-999
; CURRENT APPLICATION NUMBER: US/09/922,261
; CURRENT FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: US/09/461,697
; PRIOR FILING DATE: 1999-12-14
; NUMBER OF SEQ ID NOS: 466
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 33
; LENGTH: 93
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-922-261-33

Query Match 71.0%; Score 14.2; DB 10; Length 93;
Best Local Similarity 84.2%; Pred. No. 3.9e+02;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy       2 CCACAGGGCTCACCAGGCG 20
           |||| |||||| |||| ||
Db      28 CCACTGGGCTCCCCAGTCG 10
```

RESULT 2
US-09-813-824A-19/c
; Sequence 19, Application US/09813824A
; Patent No. US20020164595A1
; GENERAL INFORMATION:

```
;       APPLICANT: Vogelstein, Bert
;                  Kinzler, Kenneth
;                  Sherman, Michael
;       TITLE OF INVENTION: SEQUENCE SPECIFIC DNA BINDING
;                           BY P53
;       NUMBER OF SEQUENCES: 41
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Banner & Witcoff
;            STREET: 1001 G Street, NW
;            CITY: Washington
;            STATE: DC
;            COUNTRY: USA
;            ZIP: 20001
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Diskette
;            COMPUTER: IBM Compatible
;            OPERATING SYSTEM: DOS
;            SOFTWARE: FastSEQ for Windows Version 2.0
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/813,824A
;            FILING DATE: 22-Mar-2001
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 07/860,758
;            FILING DATE: 31-MAR-1992
;            APPLICATION NUMBER: 07/715,182
;            FILING DATE: 14-JUN-1991
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Kagan, Sarah A
;            REGISTRATION NUMBER: 32141
;            REFERENCE/DOCKET NUMBER: 01107.47071
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 202-508-9100
;            TELEFAX: 202-508-9299
;            TELEX: <Unknown>
;   INFORMATION FOR SEQ ID NO: 19:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 50 base pairs
;            TYPE: nucleic acid
;            STRANDEDNESS: single
;            TOPOLOGY: linear
;       SEQUENCE DESCRIPTION: SEQ ID NO: 19:
US-09-813-824A-19

  Query Match             66.0%;  Score 13.2;  DB 9;  Length 50;
  Best Local Similarity   83.3%;  Pred. No. 1.1e+03;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 ACCACAGGGCTCACCAGG 18
           ||| ||||||| |||||
Db      41 ACCCCAGGGCTGGCCAGG 24


RESULT 3
US-09-864-761-20746
; Sequence 20746, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 20746
;   LENGTH: 90
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AC008012.8
;   OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.4
;   OTHER INFORMATION: NT HIT: X99675.1, EVALUE 6.30e-01
US-09-864-761-20746

  Query Match             66.0%;  Score 13.2;  DB 10;  Length 90;
  Best Local Similarity   83.3%;  Pred. No. 1.1e+03;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 ACCACAGGGCTCACCAGG 18
           ||||||| | ||||| ||
Db      25 ACCACAGAGATCACCTGG 42


RESULT 4
US-09-761-962-37
; Sequence 37, Application US/09761962
; Patent No. US20020077285A1
; GENERAL INFORMATION:
;  APPLICANT: Memorial Sloan-Kettering Cancer Center
;  TITLE OF INVENTION: Identification and Characterization of Multiple Splice
;  TITLE OF INVENTION:  Variants of Mu-
;  TITLE OF INVENTION:  opioid Receptor (MOR-1) Gene
;  FILE REFERENCE: 830002-2000.1
;  CURRENT APPLICATION NUMBER: US/09/761,962
;  CURRENT FILING DATE:  2001-01-17
;  PRIOR APPLICATION NUMBER: 09/743,872
;  PRIOR FILING DATE: 2001-03-13
;  NUMBER OF SEQ ID NOS: 46
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 37
;   LENGTH: 24
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: primer
```

```
US-09-761-962-37

  Query Match             64.0%;  Score 12.8;  DB 10;  Length 24;
  Best Local Similarity   87.5%;  Pred. No. 1.7e+03;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       2 CCACAGGGCTCACCAG 17
           |||||  |||||||||
Db       1 CCACACTGCTCACCAG 16


RESULT 5
US-09-801-274-874
; Sequence 874, Application US/09801274
; Patent No. US20020032319A1
; GENERAL INFORMATION:
;  APPLICANT: Cargill, Michele
;  APPLICANT:  Ireland, James S.
;  APPLICANT:  Lander, Eric S.
;  TITLE OF INVENTION: HUMAN SINGLE NUCLEOTIDE POLYMORPHISMS
;  FILE REFERENCE: 2825.2009-001
;  CURRENT APPLICATION NUMBER: US/09/801,274
;  CURRENT FILING DATE:  2001-03-07
;  PRIOR APPLICATION NUMBER: US 60/187,510
;  PRIOR FILING DATE: 2000-03-07
;  PRIOR APPLICATION NUMBER: US 60/206,129
;  PRIOR FILING DATE: 2000-05-22
;  NUMBER OF SEQ ID NOS: 1802
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 874
;   LENGTH: 31
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-801-274-874

  Query Match             64.0%;  Score 12.8;  DB 10;  Length 31;
  Best Local Similarity   77.8%;  Pred. No. 1.7e+03;
  Matches   14;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 ACCACAGGGCTCACCAGG 18
           |||||| || :| |||||
Db       6 ACCACAAGGAYCCCCAGG 23


RESULT 6
US-09-783-590-11990/c
; Sequence 11990, Application US/09783590
; Patent No. US20020110850A1
; GENERAL INFORMATION:
;  APPLICANT: Dillon, Patrick J.
;  APPLICANT:  Haseltine, William A.
;  APPLICANT:  Li, Haodong
;  APPLICANT:  Rosen, Craig A.
;  APPLICANT:  Ruben, Steven M.
;  TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 16.2
;  FILE REFERENCE: PO-16.2C1
;  CURRENT APPLICATION NUMBER: US/09/783,590
;  CURRENT FILING DATE:  2000-02-15
;  PRIOR APPLICATION NUMBER: 08/420,856
;  PRIOR FILING DATE: 1995-04-12
;  PRIOR APPLICATION NUMBER: 08/346,731
;  PRIOR FILING DATE: 1994-11-21
;  NUMBER OF SEQ ID NOS: 12485
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11990
;   LENGTH: 79
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc feature
;   LOCATION: (5)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc feature
;   LOCATION: (9)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc feature
;   LOCATION: (16)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc feature
;   LOCATION: (77)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc feature
;   LOCATION: (78)
;   OTHER INFORMATION: n equals a,t,g, or c
;   NAME/KEY: misc feature
;   LOCATION: (79)
;   OTHER INFORMATION: n equals a,t,g, or c
US-09-783-590-11990

  Query Match             64.0%;  Score 12.8;  DB 10;  Length 79;
  Best Local Similarity   87.5%;  Pred. No. 1.7e+03;
  Matches   14;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       4 ACAGGGCTCACCAGGC 19
           |||||| |||||| ||
Db      40 ACAGGGTTCACCACGC 25


RESULT 7
US-09-759-136-51
; Sequence 51, Application US/09759136
; Patent No. US20010018211A1
;  GENERAL INFORMATION:
;    APPLICANT:  Chinnadurai, Govindaswamy
;    TITLE OF INVENTION:  ANTI-PROLIFERATION DOMAIN OF HUMAN Bcl-2
;    TITLE OF INVENTION:  AND DNA ENCODING THE SAME
;    NUMBER OF SEQUENCES:  51
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SUGHRUE, MION, ZINN, MACPEAK & SEAS
;      STREET:  2100 Pennsylvania Avenue, N.W.
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20037
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/759,136
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  09/274,647
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Mack, Susan J.
;      REGISTRATION NUMBER:  30,951
;      REFERENCE/DOCKET NUMBER:  A-6812
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)293-7060
;      TELEFAX:  (202)293-7860
;  INFORMATION FOR SEQ ID NO:  51:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  56 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  other nucleic acid
;      DESCRIPTION:  /desc = "Oligonucleotide Primer"
US-09-759-136-51

  Query Match             63.0%;  Score 12.6;  DB 10;  Length 56;
  Best Local Similarity   78.9%;  Pred. No. 2.1e+03;
  Matches   15;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;
```

```
Qy      1 ACCACAGGGCTCACCAGGC 19
          ||||||||   |||| |||
Db      3 ACCACAGGTGGCACCGGGC 21


RESULT 8
US-09-801-274-1365
; Sequence 1365, Application US/09801274
; Patent No. US20020032319A1
; GENERAL INFORMATION:
;  APPLICANT: Cargill, Michele
;  APPLICANT:  Ireland, James S.
;  APPLICANT:  Lander, Eric S.
;  TITLE OF INVENTION: HUMAN SINGLE NUCLEOTIDE POLYMORPHISMS
;  FILE REFERENCE: 2825.2009-001
;  CURRENT APPLICATION NUMBER: US/09/801,274
;  CURRENT FILING DATE:  2001-03-07
;  PRIOR APPLICATION NUMBER: US 60/187,510
;  PRIOR FILING DATE: 2000-03-07
;  PRIOR APPLICATION NUMBER: US 60/206,129
;  PRIOR FILING DATE: 2000-05-22
;  NUMBER OF SEQ ID NOS: 1802
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1365
;   LENGTH: 31
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-801-274-1365

  Query Match            62.0%;  Score 12.4;  DB 10;  Length 31;
  Best Local Similarity  81.2%;  Pred. No. 2.6e+03;
  Matches   13;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      3 CACAGGGCTCACCAGG 18
          ||| |||||| ||:||
Db      3 CACCGGGCTCCCCRGG 18


RESULT 9
US-09-504-231A-2132
; Sequence 2132, Application US/09504231A
; Patent No. US20020013458A1
; GENERAL INFORMATION:
;  APPLICANT: Blatt, Lawrence
;  APPLICANT:  McSwiggen, James
;  APPLICANT:  Roberts, Beth
;  APPLICANT:  Pavco, Pamela
;  APPLICANT:  Macejak, Dennis
;  TITLE OF INVENTION: ENZYMATIC NUCLEIC ACID TREATMENT OF DISEASES OR CONDITIONS RELATE
;  TITLE OF INVENTION:  HEPATITIS C VIRUS INFECTION
;  FILE REFERENCE: rpi 247/282
;  CURRENT APPLICATION NUMBER: US/09/504,231A
;  CURRENT FILING DATE:  2000-02-15
;  PRIOR APPLICATION NUMBER: 09/274,553
;  PRIOR FILING DATE: 1999-03-23
;  PRIOR APPLICATION NUMBER: 09/257,608
;  PRIOR FILING DATE: 1999-02-24
;  PRIOR APPLICATION NUMBER: 60/100,842
;  PRIOR FILING DATE: 1998-09-18
;  PRIOR APPLICATION NUMBER: 60/083,217
;  PRIOR FILING DATE: 1998-04-27
;  NUMBER OF SEQ ID NOS: 3242
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2132
;   LENGTH: 36
;   TYPE: RNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:  Enzymatic Nucleic Acid Molec
US-09-504-231A-2132

  Query Match            61.0%;  Score 12.2;  DB 10;  Length 36;
  Best Local Similarity  76.5%;  Pred. No. 3.2e+03;
  Matches   13;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      3 CACAGGGCTCACCAGGC 19
          ||||||||: |  ||||
Db      1 CACAGGGCUGAUGAGGC 17


RESULT 10
US-09-274-553D-2132
; Sequence 2132, Application US/09274553D
; Patent No. US20020082225A1
; GENERAL INFORMATION:
;  APPLICANT: Blatt, Lawrence
;  APPLICANT:  McSwiggen, James
;  APPLICANT:  Roberts, Beth
;  APPLICANT:  Pavco, Pamela
;  APPLICANT:  Macejak, Dennis
;  TITLE OF INVENTION: ENZYMATIC NUCLEIC ACID TREATMENT OF DISEASES OR CONDITIONS REL
;  TITLE OF INVENTION:  HEPATITIS C VIRUS INFECTION
;  FILE REFERENCE: rpi 247/282
;  CURRENT APPLICATION NUMBER: US/09/274,553D
;  CURRENT FILING DATE:  1999-03-23
;  PRIOR APPLICATION NUMBER: 09/257,608
;  PRIOR FILING DATE: 1999-02-24
;  PRIOR APPLICATION NUMBER: 60/100,842
;  PRIOR FILING DATE: 1998-09-18
;  PRIOR APPLICATION NUMBER: 60/083,217
;  PRIOR FILING DATE: 1998-04-27
;  NUMBER OF SEQ ID NOS: 3148
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2132
;   LENGTH: 36
;   TYPE: RNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:  Enzymatic Nucleic Acid Mo
US-09-274-553D-2132

  Query Match            61.0%;  Score 12.2;  DB 10;  Length 36;
  Best Local Similarity  76.5%;  Pred. No. 3.2e+03;
  Matches   13;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      3 CACAGGGCTCACCAGGC 19
          ||||||||: |  ||||
Db      1 CACAGGGCUGAUGAGGC 17


RESULT 11
US-09-484-704-56
; Sequence 56, Application US/09484704
; Patent No. US20020081567A1
;  GENERAL INFORMATION:
;    APPLICANT:  Henrickson, Kelly J.
;    APPLICANT:  Fan, Jiang (n.m.i.)
;    TITLE OF INVENTION:  VIRUS ASSAY METHOD
;    NUMBER OF SEQUENCES:  65
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Quarles & Brady
;      STREET:  411 East Wisconsin Avenue
;      CITY:  Milwaukee
;      STATE:  Wisconsin
;      COUNTRY:  U.S.A.
;      ZIP:  53202-4497
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/484,704
;      FILING DATE:
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
```

```
;      NAME:  Baker, Jean C.
;      REGISTRATION NUMBER:  35,433
;      REFERENCE/DOCKET NUMBER:  650053.91126
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (414) 277-5000
;      TELEFAX:  (414) 271-3552
;  INFORMATION FOR SEQ ID NO:  56:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  39 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  Oligonucleotide
US-09-484-704-56

  Query Match             61.0%;  Score 12.2;  DB 10;  Length 39;
  Best Local Similarity   82.4%;  Pred. No. 3.2e+03;
  Matches   14;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 ACCACAGGGCTCACCAG 17
           || |||| ||||| |||
Db       4 ACAACAGTGCTCAACAG 20


RESULT 12
US-09-864-761-29557/c
; Sequence 29557, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 29557
;   LENGTH: 75
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AL138956.3
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.52
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.65
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.5
;   OTHER INFORMATION: SWISSPROT HIT: P31209, EVALUE 7.40e+00
;   OTHER INFORMATION: EST_HUMAN HIT: T59472.1, EVALUE 2.20e-01
;   OTHER INFORMATION: NT HIT: Z37557.1, EVALUE 5.90e-01
US-09-864-761-29557

  Query Match             61.0%;  Score 12.2;  DB 10;  Length 75;
  Best Local Similarity   82.4%;  Pred. No. 3.3e+03;
  Matches   14;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       2 CCACAGGGCTCACCAGG 18
           ||||| |||||  ||||
Db      39 CCACATGGCTCTGCAGG 23


RESULT 13
US-09-864-761-24315/c
; Sequence 24315, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
```

```
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 24315
;   LENGTH: 97
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AL133419.11
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.9
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.1
;   OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL =   2
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL =   2
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.9
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.2
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.8
;   OTHER INFORMATION: EST_HUMAN HIT: AW956704.1, EVALUE 3.00e-01
;   OTHER INFORMATION: NT HIT: D10927.1, EVALUE 5.20e-02
US-09-864-761-24315

  Query Match             60.0%;  Score 12;  DB 10;  Length 97;
  Best Local Similarity   75.0%;  Pred. No. 4.1e+03;
  Matches   15;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy      1 ACCACAGGGCTCACCAGGCG 20
          ||||||| ||  |||| | |
Db     97 ACCACAGTGCCAACCAAGAG 78


RESULT 14
US-09-897-438B-6
; Sequence 6, Application US/09897438B
; Patent No. US20020137095A1
; GENERAL INFORMATION:
;  APPLICANT: Mikoshiba, Katsuhiko
;  APPLICANT:  Tate, Naoko
;  TITLE OF INVENTION: REELIN PROTEIN CR-50 EPITOPE REGION
;  FILE REFERENCE: 04853-0076-00000
;  CURRENT APPLICATION NUMBER: US/09/897,438B
;  CURRENT FILING DATE:  2001-07-03
;  PRIOR APPLICATION NUMBER: JP 2000-202801
;  PRIOR FILING DATE: 2000-07-04
;  NUMBER OF SEQ ID NOS: 11
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
;   LENGTH: 22
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: synthetic
;   OTHER INFORMATION: primer for PCR
US-09-897-438B-6

  Query Match             59.0%;  Score 11.8;  DB 10;  Length 22;
  Best Local Similarity   86.7%;  Pred. No. 4.8e+03;
  Matches   13;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      5 CAGGGCTCACCAGGC 19
          ||||||  || |||||
Db      1 CAGGGCCCAGCAGGC 15


RESULT 15
US-09-801-274-915
; Sequence 915, Application US/09801274
; Patent No. US20020032319A1
; GENERAL INFORMATION:
;  APPLICANT: Cargill, Michele
;  APPLICANT:  Ireland, James S.
;  APPLICANT:  Lander, Eric S.
;  TITLE OF INVENTION: HUMAN SINGLE NUCLEOTIDE POLYMORPHISMS
;  FILE REFERENCE: 2825.2009-001
;  CURRENT APPLICATION NUMBER: US/09/801,274
;  CURRENT FILING DATE:  2001-03-07
;  PRIOR APPLICATION NUMBER: US 60/187,510
;  PRIOR FILING DATE: 2000-03-07
;  PRIOR APPLICATION NUMBER: US 60/206,129
;  PRIOR FILING DATE: 2000-05-22
;  NUMBER OF SEQ ID NOS: 1802
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 915
;   LENGTH: 31
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-801-274-915

  Query Match             59.0%;  Score 11.8;  DB 10;  Length 31;
  Best Local Similarity   76.5%;  Pred. No. 4.9e+03;
  Matches   13;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy      2 CCACAGGGCTCACCAGG 18
          || |||||  |||:|||
Db      3 CCGCAGGGTCCACYAGG 19


Search completed: November 23, 2002, 07:10:45
Job time : 17.8 secs
```

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 26, 2002, 04:16:10 ; Search time 798.5 Seconds
(without alignments)
405.648 Million cell updates/sec

Title: US-09-296-264-30
Perfect score: 20
Sequence: 1 accacagggctcaccaggcg 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 357874

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	14.2	71.0	72	10	AW245565	AW245565 2822910.5
c	2	13.8	69.0	80	9	AA389474	AA389474 mp25g06.r
	3	13.6	68.0	71	17	AZ871187	AZ871187 2M0183L24
c	4	13.6	68.0	73	9	AA003808	AA003808 mg78a03.r
	5	13.6	68.0	81	14	H80149	H80149 yu11g01.s1
c	6	13.6	68.0	85	12	BG314931	BG314931 OP2.0.244
	7	13.6	68.0	86	9	AA883180	AA883180 am18a11.s
	8	13.6	68.0	96	9	AJ282942	AJ282942 4A3A-P2B1
	9	13.6	68.0	99	9	AA196710	AA196710 zq59e09.r
	10	13.4	67.0	19	9	AI431460	AI431460 th40c01.x
c	11	13.4	67.0	43	9	AA506141	AA506141 ni17f03.s
c	12	13.2	66.0	24	17	AZ810559	AZ810559 2M0076I07
	13	13.2	66.0	48	17	AZ818575	AZ818575 2M0088M10
c	14	13.2	66.0	49	9	AI889232	AI889232 wm36b09.x
	15	13.2	66.0	66	17	AZ779239	AZ779239 2M0015G22
c	16	13.2	66.0	67	9	AA976767	AA976767 oq09d07.s
	17	13.2	66.0	68	9	AA000116	AA000116 mg33a05.r
c	18	13.2	66.0	88	9	AI527132	AI527132 uj49g10.x
c	19	13.2	66.0	90	9	AA503993	AA503993 nh39e01.s
	20	13.2	66.0	91	17	AZ435316	AZ435316 1M0222C18
	21	13	65.0	33	13	BJ079504	BJ079504 BJ079504
c	22	12.8	64.0	32	17	AZ471070	AZ471070 1M0285A02
c	23	12.8	64.0	41	17	AZ789628	AZ789628 2M0037A09
c	24	12.8	64.0	50	9	AU105202	AU105202 AU105202
	25	12.8	64.0	50	9	AU106633	AU106633 AU106633
	26	12.8	64.0	50	17	BH797330	BH797330 1008087F1
c	27	12.8	64.0	58	17	AZ369075	AZ369075 1M0119A05
	28	12.8	64.0	72	17	AZ817739	AZ817739 2M0087E20
	29	12.8	64.0	76	9	AA692294	AA692294 vt20d01.r
	30	12.8	64.0	90	17	AZ943059	AZ943059 2M0203M03
	31	12.8	64.0	93	14	R97537	R97537 yq56f07.s1
	32	12.8	64.0	94	9	AI374859	AI374859 ta62d10.x
c	33	12.8	64.0	95	9	AA506784	AA506784 EST022 Hu
c	34	12.6	63.0	27	17	AZ579571	AZ579571 1M0367M03
	35	12.6	63.0	42	9	AL634179	AL634179 AL634179
	36	12.6	63.0	53	9	AA947039	AA947039 oq57c10.s
	37	12.6	63.0	53	9	AA564013	AA564013 nk24b09.s
	38	12.6	63.0	62	17	AZ920920	AZ920920 1006022A1
	39	12.6	63.0	64	10	AW626721	AW626721 SWOvAFCAP
	40	12.6	63.0	64	13	BI097425	BI097425 SWOv3MCAM
	41	12.6	63.0	70	9	AI196928	AI196928 ui68a09.x
	42	12.6	63.0	75	17	AZ920782	AZ920782 1006022E1
	43	12.6	63.0	77	14	R09359	R09359 yf26g01.r1
	44	12.6	63.0	78	12	BG328037	BG328037 602427674
c	45	12.6	63.0	78	13	BI851925	BI851925 603378084

ALIGNMENTS

RESULT 1
AW245565/c
LOCUS AW245565 72 bp mRNA linear EST 07-JAN-2000
DEFINITION 2822910.5prime NIH_MGC_7 Homo sapiens cDNA clone IMAGE:2822910 5', mRNA sequence.
ACCESSION AW245565
VERSION AW245565.1 GI:6588558
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 72)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Other_ESTs: 2822910.3prime
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTP cDNA Library Preparation: Ling Hong/Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Berkeley MGC sequencing project Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html Base Calling / Quality Scores: PHRED from University of Washington Genome Center. Vector Trimming: cross_match from University of Washington Genome Center PHRAP suite. Poly-T Identification: patMatch.pl from Berkeley Drosophila Genome Project. University of Washington Genome Center:

```
          http://www.genome.washington.edu Low Quality Sequence: 0 contiguous
          PHRED high quality bases following vector sequence. Very Low
          Quality Sequence: Trace file contained 72 contiguous distinct peaks
          following vector sequence.
          Plate: LLCM10  row: K  column: 7.
FEATURES             Location/Qualifiers
     source          1. .72
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2822910"
                     /clone_lib="NIH_MGC_7"
                     /tissue_type="small cell carcinoma"
                     /cell_line="MGC3"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: lung; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming.  Directionally
                     cloned into EcoRI/XhoI sites using the following 5'
                     adaptor: GGCACGAG(G).  Size-selected >500bp for average
                     insert size 1.8kb.  Library constructed by Ling Hong in
                     the laboratory of Gerald M. Rubin (University of
                     California, Berkeley) using  ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT       14 a     21 c     30 g      7 t
ORIGIN

  Query Match             71.0%;  Score 14.2;  DB 10;  Length 72;
  Best Local Similarity   84.2%;  Pred. No. 2.5e+04;
  Matches   16;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 ACCACAGGGCTCACCAGGC 19
           ||| |||||||||| | |||
Db      46 ACCTCAGGGCTCAGCCGGC 28


RESULT 2
AA389474/c
LOCUS       AA389474                 80 bp    mRNA    linear   EST 23-APR-1997
DEFINITION  mp25g06.r1 Life Tech mouse embryo 8 5dpc 10664019 Mus musculus cDNA
            clone IMAGE:570298 5' similar to gb:X89650 M.musculus mRNA for Rab7
            protein (MOUSE);, mRNA sequence.
ACCESSION   AA389474
VERSION     AA389474.1  GI:2042429
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 80)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:344946
            Trace considered overall poor quality
            Seq primer: -28m13 rev1 ET from Amersham
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .80
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:570298"
                     /clone_lib="Life Tech mouse embryo 8 5dpc 10664019"
                     /tissue_type="embryo"
                     /dev_stage="8.5dpc embryos"
                     /lab_host="DH10B"
                     /note="Organ: whole embryo; Vector: pCMV-SPORT2; Site_1:
                     SalI; Site_2: NotI; Cloned unidirectionally.  Primer:
                     Oligo dT.  8.5dpc embryos. pCMV-SPORT2 vector."
BASE COUNT       10 a     25 c     28 g     17 t
ORIGIN

  Query Match             69.0%;  Score 13.8;  DB 9;  Length 80;
  Best Local Similarity   88.2%;  Pred. No. 3.8e+04;
  Matches   15;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       3 CACAGGGCTCACCAGGC 19
           |||||||| ||| ||||
Db      63 CACAGGGCCCACGAGGC 47


RESULT 3
AZ871187
LOCUS       AZ871187                 71 bp    DNA     linear   GSS 21-FEB-2001
DEFINITION  2M0183L24R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC2M0183L24 R, DNA sequence.
ACCESSION   AZ871187
VERSION     AZ871187.1  GI:13077187
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 71)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0183  row: L  column: 24
            Seq primer: CACACAGGAAACAGCTATGACC
            Class: plasmid ends
            High quality sequence stop: 71.
FEATURES             Location/Qualifiers
     source          1. .71
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC2M0183L24"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
```

electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptored mouse DNA was annealed to adaptored vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 18 a 11 c 21 g 21 t
ORIGIN

Query Match 68.0%; Score 13.6; DB 17; Length 71;
Best Local Similarity 80.0%; Pred. No. 4.4e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 ACCACAGGGCTCACCAGGCG 20
||||||| |||||||| |
Db 19 ACCACAGACATCACCAGGTG 38

RESULT 4
AA003808/c
LOCUS AA003808 73 bp mRNA linear EST 23-JUL-1996
DEFINITION mg78a03.r1 Soares mouse embryo NbME13.5 14.5 Mus musculus cDNA clone IMAGE:439084 5' similar to PIR:S19586 S19586 N-methyl-D-aspartate receptor glutamate-binding chain - rat ;, mRNA sequence.
ACCESSION AA003808
VERSION AA003808.1 GI:1447283
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 73)
AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.
TITLE The WashU-HHMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Marra M/Mouse EST Project
WashU-HHMI Mouse EST Project
Washington University School of MedicineP
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:264420
Trace considered overall poor quality
Possible reversed clone: similarity on wrong strand
Seq primer: mob.REGA+ET
High quality sequence stop: 1.
FEATURES Location/Qualifiers
source 1. .73
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:439084"
/clone_lib="Soares mouse embryo NbME13.5 14.5"
/sex="unknown"
/tissue_type="embryo"
/dev_stage="13.5-14.5dpc total fetus"
/lab_host="DH10B"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGGAGCGGCCGCGGAAATTTTTTTTTTTTTTTTTTTTT T 3'], on equal amounts of mRNA from 2 13.5dpc and 2 14.5dpc embryos [total RNA provided by Minoru Ko, Wayne State Univ., from 2]; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library went through one round of normalization, and was constructed by Bento Soares and M.Fatima Bonaldo. "

BASE COUNT 6 a 22 c 22 g 23 t
ORIGIN

Query Match 68.0%; Score 13.6; DB 9; Length 73;
Best Local Similarity 80.0%; Pred. No. 4.5e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 ACCACAGGGCTCACCAGGCG 20
|||||| |||||||| |
Db 51 ACCACAACACTCACCAGGAG 32

RESULT 5
H80149
LOCUS H80149 81 bp mRNA linear EST 09-NOV-1995
DEFINITION yu11g01.s1 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone IMAGE:233520 3' similar to gb:M75126 HEXOKINASE, TYPE I (HUMAN);, mRNA sequence.
ACCESSION H80149
VERSION H80149.1 GI:1058238
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 81)
AUTHORS Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiapelli,B., Chissoe,S., Dietrich,N., DuBuque,T., Favello,A., Gish,W., Hawkins,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N., Mardis,E., Moore,B., Morris,M., Parsons,J., Prange,C., Rifkin,L., Rohlfing,T., Schellenberg,K., Soares,M.B., Tan,F., Thierry-Meg,J., Trevaskis,E., Underwood,K., Wohldmann,P., Waterston,R., Wilson,R. and Marra,M.
TITLE Generation and analysis of 280,000 human expressed sequence tags
JOURNAL Genome Res. 6 (9), 807-828 (1996)
MEDLINE 97044478
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Insert Size: 703
High quality sequence starts: 1
High quality sequence stops: 1
Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Trace considered overall poor quality
Insert Length: 703 Std Error: 0.00
Seq primer: Promega -21m13
High quality sequence stop: 1.
FEATURES Location/Qualifiers
source 1. .81
/organism="Homo sapiens"
/db_xref="GDB:3786697"
/db_xref="taxon:9606"
/clone="IMAGE:233520"
/clone_lib="Soares fetal liver spleen 1NFLS"
/sex="male"
/dev_stage="20 week-post conception fetus"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia) with a modified polylinker; Site_1: Pac I; Site_2: Eco RI; 1st strand cDNA was primed with a Pac I - oligo(dT) primer [5' AACTGGAAGAATTAATTAAAGATCTTTTTTTTTTTTTTTTTT 3'], double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Pac I and cloned into the Pac I

and Eco RI sites of the modified pT7T3 vector. Library went through one round of normalization. Library constructed by Bento Soares and M.Fatima Bonaldo."
BASE COUNT 15 a 17 c 20 g 24 t 5 others
ORIGIN

Query Match 68.0%; Score 13.6; DB 14; Length 81;
Best Local Similarity 80.0%; Pred. No. 4.6e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 ACCACAGGGCTCACCAGGCG 20
|| ||| ||||||| | |||
Db 21 ACGACACGGCTCACAAAGCG 40

RESULT 6
BG314931/c
LOCUS BG314931 85 bp mRNA linear EST 14-MAR-2002
DEFINITION OP2.0.244 Human THP1 cell line library Homo sapiens cDNA, mRNA sequence.
ACCESSION BG314931
VERSION BG314931.1 GI:18997795
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 85)
AUTHORS Andersson,T., Borang,S., Larsson,M., Thelin,A., Ekstrand-Hammarsrom,B., Wirta,V., Wenborg,A., Lundeberg,J. and Odeberg,J.
TITLE Identification of candidate genes in atherosclerosis - Virtual chip analysis in RDA based transcript profiling of monocyte/macrophage response to oxidised LDL
JOURNAL Unpublished (2001)
COMMENT Contact: Andersson Tove
Department of Biotechnology
KTH
Teknikringen 34, plan 6, 100 44 Stockholm, Sweden
Tel: +46 8 790 71 29
Fax: +46 8 245452
Email: tove@biochem.kth.se
POLYA=No.
FEATURES Location/Qualifiers
source 1. .85
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="Human THP1 cell line library"
/cell_line="THP1"
/note="Vector: pRIT28; Site_1: BamHI; Site_2: BamHI; Shotgun cloning of RDA difference products. Macrophage and foamcell libraries were submitted to successive rounds of subtractive hybridisations generating populations of gene fragments that are differentially expressed in macrophage to foam cell formation."
BASE COUNT 15 a 17 c 37 g 16 t
ORIGIN

Query Match 68.0%; Score 13.6; DB 12; Length 85;
Best Local Similarity 80.0%; Pred. No. 4.7e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 ACCACAGGGCTCACCAGGCG 20
| ||||||| || |||| ||
Db 24 AGCACAGGGTTCTCCAGCCG 5

RESULT 7
AA883180
LOCUS AA883180 86 bp mRNA linear EST 04-JAN-1999
DEFINITION am18a11.s1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone IMAGE:1467164 3' similar to gb:M75126 HEXOKINASE, TYPE I (HUMAN);, mRNA sequence.
ACCESSION AA883180
VERSION AA883180.1 GI:2994656
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 86)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 552 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham.
FEATURES Location/Qualifiers
source 1. .86
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1467164"
/clone_lib="Soares_NFL_T_GBC_S1"
/lab_host="DH10B"
/note="Organ: pooled; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; Equal amounts of plasmid DNA from three normalized libraries (fetal lung NbHL19W, testis NHT, and B-cell NCI_CGAP_GCB1) were mixed, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 3 libraries. The pools consisted of I.M.A.G.E. clones 297480-302087, 682632-687239, 726408-728711, and 729096-731399. Subtraction by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT 22 a 14 c 24 g 26 t
ORIGIN

Query Match 68.0%; Score 13.6; DB 9; Length 86;
Best Local Similarity 80.0%; Pred. No. 4.8e+04;
Matches 16; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 ACCACAGGGCTCACCAGGCG 20
|| ||| ||||||| | |||
Db 27 ACGACACGGCTCACAAAGCG 46

RESULT 8
AJ282942
LOCUS AJ282942 96 bp mRNA linear EST 30-JUN-2000
DEFINITION 4A3A-P2B10-R Anopheles gambiae immune competent 4A3A Anopheles gambiae cDNA clone 4A3A-P2B10, mRNA sequence.
ACCESSION AJ282942
VERSION AJ282942.1 GI:6930821
KEYWORDS EST.
SOURCE African malaria mosquito.
ORGANISM Anopheles gambiae
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anopheles.
REFERENCE 1 (bases 1 to 96)
AUTHORS Dimopoulos,G., Casavant,T.L., Chang,S., Scheetz,T., Roberts,C., Donohue,M., Schultz,J., Benes,V., Bork,P., Ansorge,W., Soares,M.B. and Kafatos,F.C.
TITLE Anopheles gambiae pilot gene discovery project: identification of mosquito innate immunity genes from expressed sequence tags generated from immune-competent cell lines
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (12), 6619-6624 (2000)
MEDLINE 20300950
COMMENT Contact: Dimopoulos G
Fotis C. Kafatos laboratory
European Molecular Biology Laboratory

```
          Meyerhofstrasse 1, 69117 Heidelberg, Germany.
FEATURES             Location/Qualifiers
     source          1. .96
                     /organism="Anopheles gambiae"
                     /strain="4A r/r"
                     /db_xref="taxon:7165"
                     /clone="4A3A-P2B10"
                     /clone_lib="Anopheles gambiae immune competent 4A3A"
                     /cell_line="immune competent 4A3A"
                     /lab_host="E. coli DH10B"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: EcoRI; Site_2: NotI; sequenced from
                     forward priming site which reads from the 3' end of the
                     cDNA. The 4A3A is a directionally cloned and normalized
                     cDNA library that was constructed from the 4A3A cell line
                     oligo-T primed cDNA according to: Bonaldo, Lennon & Soares
                     (1996) : Normalization and Subtraction: Two approaches To
                     Facilitate Gene Discovery, Genome Research 6, 791-806."
BASE COUNT       17 a     32 c     32 g     15 t
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 9;  Length 96;
  Best Local Similarity   80.0%;  Pred. No. 5e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 ACCACAGGGCTCACCAGGCG 20
          |||| | |||| ||| ||||
Db     74 ACCAGAAGGCTAACCTGGCG 93


RESULT 9
AA196710
LOCUS       AA196710                 99 bp    mRNA    linear   EST 22-JAN-1997
DEFINITION  zq59e09.r1 Stratagene neuroepithelium (#937231) Homo sapiens cDNA
            clone IMAGE:645928 5' similar to gb:M75126 HEXOKINASE, TYPE I
            (HUMAN);, mRNA sequence.
ACCESSION   AA196710
VERSION     AA196710.1  GI:1792301
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 99)
  AUTHORS   Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiapelli,B.,
            Chissoe,S., Dietrich,N., DuBuque,T., Favello,A., Gish,W., Hawkins
            ,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N., Mardis,E., Moore
            ,B., Morris,M., Parsons,J., Prange,C., Rifkin,L., Rohlfing,T.,
            Schellenberg,K., Soares,M.B., Tan,F., Thierry-Meg,J., Trevaskis,E.,
            Underwood,K., Wohldmann,P., Waterston,R., Wilson,R. and Marra,M.
  TITLE     Generation and analysis of 280,000 human expressed sequence tags
  JOURNAL   Genome Res. 6 (9), 807-828 (1996)
  MEDLINE   97044478
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: -28M13 rev2 from Amersham
            High quality sequence stop: 29.
FEATURES             Location/Qualifiers
     source          1. .99
                     /organism="Homo sapiens"
                     /db_xref="GDB:5216179"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:645928"
                     /clone_lib="Stratagene neuroepithelium (#937231)"
                     /dev_stage="Ntera-2/RA neuroepithelial cells"
                     /lab_host="SOLR (kanamycin resistant)"
                     /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
                     XhoI; Cloned unidirectionally.  Primer: Oligo dT. NT2
                     cells (Ntera-2/c1.D1) induced with Retinoic Acid for 24
                     hours. Average insert size: 1.5 kb; Uni-ZAP XR Vector; ~5'
                     adaptor sequence: 5' GAATTCGGCACGAG 3' ~3' adaptor
                     sequence: 5' CTCGAGTTTTTTTTTTTTTTTTT 3'"
BASE COUNT       25 a     17 c     24 g     29 t      4 others
ORIGIN

  Query Match             68.0%;  Score 13.6;  DB 9;  Length 99;
  Best Local Similarity   80.0%;  Pred. No. 5e+04;
  Matches   16;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      1 ACCACAGGGCTCACCAGGCG 20
          || ||| ||||||| | |||
Db     26 ACGACACGGCTCACAAAGCG 45


RESULT 10
AI431460
LOCUS       AI431460                 19 bp    mRNA    linear   EST 13-APR-1999
DEFINITION  th40c01.x1 NCI_CGAP_Lym12 Homo sapiens cDNA clone IMAGE:2120736 3'
            similar to TR:Q04117 Q04117 SALIVARY PROLINE-RICH PROTEIN RP4
            PRECURSOR. ;contains element MSR1 repetitive element ;, mRNA
            sequence.
ACCESSION   AI431460
VERSION     AI431460.1  GI:4303341
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 19)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Life Technologies catalog #: 11547-015
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html

            Trace considered overall poor quality
            Insert Length: 653   Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .19
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2120736"
                     /clone_lib="NCI_CGAP_Lym12"
                     /tissue_type="lymphoma, follicular mixed small and large
                     cell"
                     /lab_host="DH10B"
                     /note="Organ: lymph node; Vector: pCMV-SPORT6; Site_1:
                     SalI; Site_2: NotI; Cloned unidirectionally.  Primer:
                     Oligo dT.  Average insert size 1.25 kb. Life Technologies
                     catalog #: 11547-015"
BASE COUNT        5 a     11 c      3 g      0 t
ORIGIN

  Query Match             67.0%;  Score 13.4;  DB 9;  Length 19;
  Best Local Similarity   93.3%;  Pred. No. 3.2e+04;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 ACCACAGGGCTCACC 15
          |||||||||| ||||
Db      1 ACCACAGGGCCCACC 15
```

```
RESULT 11
AA506141/c
LOCUS       AA506141                 43 bp    mRNA    linear   EST 20-AUG-1997
DEFINITION  ni17f03.s1 NCI_CGAP_Co4 Homo sapiens cDNA clone IMAGE:968285 3'
            similar to SW:ATP6_HUMAN P00846 ATP SYNTHASE A CHAIN ;, mRNA
            sequence.
ACCESSION   AA506141
VERSION     AA506141.1  GI:2242381
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 43)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Elias Campo, M.D., Michael R. Emmert-Buck, M.D.
            , Ph.D.
            cDNA Library Preparation: M. Bento Soares, Ph.D.
            cDNA Library Arrayed by: Greg Lennon, Ph.D.
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html

            Trace considered overall poor quality
            Insert Length: 665   Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .43
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:968285"
                     /clone_lib="NCI_CGAP_Co4"
                     /sex="pooled"
                     /tissue_type="colon"
                     /lab_host="DH10B"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; 1st strand cDNA was prepared from pooled colon
                     tumor tissue, and was then primed with a Not I - oligo(dT)
                     primer. Double-stranded cDNA was ligated to Eco RI
                     adaptors (Pharmacia), digested with Not I and cloned into
                     the Not I  and Eco RI sites of the modified pT7T3 vector.
                     This library is not normalized. Library constructed by
                     Bento Soares and M. Fatima Bonaldo. "
BASE COUNT        8 a      4 c     22 g      9 t
ORIGIN

  Query Match             67.0%;  Score 13.4;  DB 9;  Length 43;
  Best Local Similarity   93.3%;  Pred. No. 4.4e+04;
  Matches   14;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 ACCACAGGGCTCACC 15
           ||||||||||| ||||
Db      32 ACCACAGGGCACACC 18


RESULT 12
AZ810559/c
LOCUS       AZ810559                 24 bp    DNA     linear   GSS 20-FEB-2001
DEFINITION  2M0076I07F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC2M0076I07 F, DNA sequence.
ACCESSION   AZ810559
VERSION     AZ810559.1  GI:12977929
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 24)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0076  row: I  column: 07
            Seq primer: CGTTGTAAAACGACGGCCAGT
            Class: plasmid ends
            High quality sequence stop: 24.
FEATURES             Location/Qualifiers
     source          1. .24
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC2M0076I07"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
BASE COUNT        4 a      4 c      9 g      7 t
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 17;  Length 24;
  Best Local Similarity   83.3%;  Pred. No. 4.3e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       3 CACAGGGCTCACCAGGCG 20
           ||||| |||| ||| |||
Db      24 CACAGAGCTCCCCAAGCG 7


RESULT 13
AZ818575
LOCUS       AZ818575                 48 bp    DNA     linear   GSS 20-FEB-2001
DEFINITION  2M0088M10R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC2M0088M10 R, DNA sequence.
ACCESSION   AZ818575
VERSION     AZ818575.1  GI:12988483
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

REFERENCE 1 (bases 1 to 48)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D.,Weiss,R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0088 row: M column: 10
Seq primer: CACACAGGAAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 48.

FEATURES Location/Qualifiers
source 1. .48
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC2M0088M10"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptored DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptored mouse DNA was annealed to adaptored vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 14 a 20 c 8 g 6 t
ORIGIN

Query Match 66.0%; Score 13.2; DB 17; Length 48;
Best Local Similarity 83.3%; Pred. No. 5.6e+04;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 CCACAGGGCTCACCAGGC 19
||||| | |||||||| |
Db 11 CCACAAGTCTCACCAGCC 28

RESULT 14
AI889232/c
LOCUS AI889232 49 bp mRNA linear EST 07-MAR-2000
DEFINITION wm36b09.x1 NCI_CGAP_Ut4 Homo sapiens cDNA clone IMAGE:2438009 3' similar to SW:RL3_HUMAN P39023 60S RIBOSOMAL PROTEIN L3 ;, mRNA sequence.
ACCESSION AI889232
VERSION AI889232.1 GI:5594396
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 49)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 1481 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 1.

FEATURES Location/Qualifiers
source 1. .49
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2438009"
/clone_lib="NCI_CGAP_Ut4"
/tissue_type="serous papillary carcinoma, high grade, 2 pooled tumors"
/lab_host="DH10B"
/note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: SalI; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.48 kb. Life Technologies catalog #: 11542-016"

BASE COUNT 6 a 16 c 10 g 16 t 1 others
ORIGIN

Query Match 66.0%; Score 13.2; DB 9; Length 49;
Best Local Similarity 83.3%; Pred. No. 5.6e+04;
Matches 15; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 3 CACAGGGCTCACCAGGCG 20
|| |||| ||||||| ||
Db 38 CAAAGGGGTCACCAGTCG 21

RESULT 15
AZ779239
LOCUS AZ779239 66 bp DNA linear GSS 16-FEB-2001
DEFINITION 2M0015G22F Mouse 10kb plasmid UUGC1M library Mus musculus genomic clone UUGC2M0015G22 F, DNA sequence.
ACCESSION AZ779239
VERSION AZ779239.1 GI:12909691
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 66)
AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D.,Weiss,R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu

```
          Insert Length: 10000   Std Error: 0.00
          Plate: 0015  row: G  column: 22
          Seq primer: CGTTGTAAAACGACGGCCAGT
          Class: plasmid ends
          High quality sequence stop: 66.
FEATURES             Location/Qualifiers
     source          1. .66
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC2M0015G22"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
BASE COUNT       13 a     21 c     18 g     14 t
ORIGIN

  Query Match             66.0%;  Score 13.2;  DB 17;  Length 66;
  Best Local Similarity   83.3%;  Pred. No. 6.3e+04;
  Matches   15;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 ACCACAGGGCTCACCAGG 18
           ||||||||||  | ||||
Db       4 ACCACAGGGCAGAACAGG 21


Search completed: November 26, 2002, 17:58:40
Job time : 808.5 secs
```

